The Dissertation Committee for Lindsey Elizabeth Gulden certifies that this is the approved version of the following dissertation:

Quantification of the confidence that can be placed in land-surface model predictions: Applications to vegetation and hydrologic processes.

Committee:

Zong-Liang Yang, Supervisor

David T. Allen

Robert E. Dickinson

Guo-Yue Niu

Clark R. Wilson

Quantification of the confidence that can be placed in land-surface model predictions: Applications to vegetation and hydrologic processes

by

Lindsey Elizabeth Gulden, A.B.

Dissertation

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

Doctor of Philosophy

The University of Texas at Austin

August, 2009

Dedication

For my great grandfather, who instilled in me the value of education; for my mother, who taught me to think for myself; and for my father, who cultivated my curiosity.

Acknowledgements

This work was funded by the National Science Foundation, the San Antonio Area Foundation, the National Aeronautics and Space Administration (NASA), the Environmental Protection Agency, and the Jackson School of Geosciences. The National Center for Atmospheric Research, NASA, and the Texas Advanced Computing Center provided computing resources.

The coauthors of the papers presented here made significant contributions to both the substance and style of the published research. I thank them for their insight, time, energy, and resources.

By pointing me toward the geosciences, Dr. John Southard and Wynne Miller enabled me to benefit from the help and guidance of a large cadre of individuals. Dr. Liang Yang, my supervisor, was the major force in my scientific development; his direction and support enabled me to complete the research presented here. At the beginning of my doctoral studies, Dr. Charles Jackson framed my understanding of how models are used and continued to advise and challenge me with humility, insight, and integrity throughout my tenure as a student. My capacity for scientific thought and reasoning benefitted from the time and insight of numerous researchers including Dr. David Allen, Dr. Jay Banner, Dr. Robert Dickinson, Dr. Ying Fan, Dr. Dave Gochis, Dr. Martin Jackson, Dr. Guo-Yue Niu, Dr. Roy Rasmussen, Dr. Alan Robock, Dr. Matthew Rodell, Dr. Bridget Scanlon, Dr. Jack Sharp, Dr. Thorsten Wagener, Dr. Christine Wiedinmyer, and Dr. Clark Wilson. With great competence and good humor, Philip Guerrero helped me navigate through the University red tape. Xiaoyan Jiang, Dr. Marla Lowrey, and Hua Su were sounding boards for countless discussions. Enrique Rosero continuously challenged me to take a second look at what I thought I knew.

Quantification of the confidence that can be placed in land-surface model predictions: Applications to vegetation and hydrologic processes

Publication No._____

Lindsey Elizabeth Gulden, Ph.D.

The University of Texas at Austin, 2009

Supervisor: Zong-Liang Yang

Abstract

The research presented here informs the confidence that can be placed in the simulations of land-surface models (LSMs).

After introducing a method for simplifying a complex, heterogeneous land-cover dataset for use in LSMs, I show that LSMs can realistically represent the spatial distribution of heterogeneous land-cover processes (e.g., biogenic emission of volatile organic compounds) in Texas. LSM-derived estimates of biogenic emissions are sensitive (varying up to a factor of 3) to land-cover data, which is not well constrained by observations. Simulated emissions are most sensitive to land-cover data in eastern and central Texas, where tropospheric ozone pollution is a concern. I further demonstrate that interannual variation in leaf mass is at least as important to variation in biogenic emissions as is interannual variation in shortwave radiation and temperature. Model

estimates show that more-humid regions with less year-to-year variation in precipitation have lower year-to-year variation in biogenic emissions: as modeled mean emissions increase, their mean-normalized standard deviation decreases.

I evaluate three parameterizations of subsurface hydrology in LSMs (with (1) a shallow, 10-layer soil; (2) a deeper, many-layered soil; and (3) a lumped aquifer model) under increasing parameter uncertainty. When given their optimal parameter sets, all three versions perform equivalently well when simulating monthly change in terrestrial water storage. The most conceptually realistic model is least sensitive to errant parameter values. However, even when using the most conceptually realistic model, parameter interaction ensures that knowing ranges for individual parameters is insufficient to guarantee realistic simulation.

LSMs are often developed and evaluated at data-rich sites but are then applied in regions where data are sparse or unavailable. I present a framework for model evaluation that explicitly acknowledges perennial sources of uncertainty in LSM simulations (e.g., parameter uncertainty, meteorological forcing-data uncertainty, evaluation-data uncertainty) and that evaluates LSMs in a way that is consistent with models' typical application. The model performance score quantifies the likelihood that a representative ensemble of model performance will bracket observations with high skill and low spread. The robustness score quantifies the sensitivity of model performance to parameter error or data error. The fitness score ranks models' suitability for broad application.

Table of Contents

List of Tables

List of Figures

Chapter 1: Introduction

The chapters of this dissertation reflect the evolution of my understanding of the merits and limitations of land-surface models and my subsequent thinking regarding ways in which the land-surface modeling community can improve the evaluation and development of land-surface models given the models' inherent predictive uncertainty. All research presented here in some way informs the confidence that can be placed in the simulations of LSMs.

1.1. LAND-SURFACE MODELS (LSMS)

LSMs represent physical processes near the Earth's surface. First invented as a means for generating more realistic lower boundary conditions for numerical models of the atmosphere (Manabe, 1969), LSMs simulate the fluxes of energy, water, and momentum between the surface and the atmosphere and within the near-surface soil, vegetation, and snow (Pitman et al., 2003). Used in weather and climate prediction, LSMs are integral components of weather and climate research. They are starting to be employed as stand-alone models to inform water resources and agricultural applications (e.g., Mo, 2008).

Like the simulations of all models of complex natural systems, the simulations of LSMs inherit uncertainty from the data used to drive them, from their simplified representations of physical processes, and from the values of their model parameters (Wagener and Gupta, 2005). LSMs are at once both insufficiently complex and too complex. Processes that occur at the molecular level (e.g., evaporation, photosynthesis)

influence regional patterns (e.g., Shuttleworth, 2007). We have yet to understand the intricacies and drivers of such molecular-level processes sufficiently well to model them at coarse resolutions and across sites. Furthermore, even though LSMs are orders of magnitude less complex than nature, current-generation models are poorly constrained (Hogue et al., 2006; Rosero et al., 2009) because of a dearth of data (Lyon et al., 2008) with which to properly evaluate their process representations. Even with sufficient understanding and adequate validation data, we still lack sufficient computational power for representing the intricate, coupled physical processes that shape land-surface−atmosphere interactions (Kumar et al., 2006). Yet, in spite of these limitations, today's LSMs perform remarkably well (Mitchell et al., 2004; Stockli et al., 2008). LSMs are arguably our best currently available tool to apply the scientific method to hypotheses about how the land surface functions within the Earth system (Pitman, 2003).

1.2. OVERVIEW OF WORK PRESENTED HERE

I began my dissertation research with a straightforward task. I was to simplify a complicated dataset that represented the land cover of Texas as a fine-scale mosaic of more than 600 plant species. The simplified dataset was to be used within an LSM to help quantify the roles of climate change and land cover change in shaping variation in the emission of volatile organic compounds from vegetation. While developing the bottom-up classification method that I used to transform the species-based land-cover dataset (Gulden and Yang, 2006; see Chapter 2), I gained insight into how modelers think about representing complex systems, how simplifying assumptions necessarily add uncertainty to simulations, and how input data can contribute significant uncertainty to model-based

estimates. Although simple in concept, the problem that I first addressed – How can a heterogeneous environment be represented in simplified form while still maintaining the essence of the system? – is an analog for the question that underlies all of environmental modeling. Modelers attempt to condense a complex system to its essential features and processes, hoping that their simplifying assumptions do not compromise what they believe to be the system's most important functional modes.

Following my increased awareness of input-data uncertainty, I used a simplified sensitivity analysis to explore the extent to which model output (in this case, simulated biogenic emissions) depends on the uncertainty contained within one of many input data sources (Gulden et al., 2008a; see Chapter 3). The work showed that BVOC estimates are very sensitive to land-cover data, which are not well constrained by observations.

To make model simulations computationally tractable and verifiable using available observations, some processes that occur in nature must either be omitted from models or loosely parameterized according to the model developer's scientific understanding. These necessary simplifications create process uncertainty that unavoidably contributes to uncertainty in model simulations. In the case of biogenic emissions, a representation of interannual variation in leaf mass (a result of year-to-year changes in environmental conditions) had been traditionally omitted from LSMs, the current generation of which specify a fixed annual cycle for vegetation growth and senescence. Year-to-year change in biomass could be important, especially to processes that depend on the quantity of biomass on the landscape (e.g., the emission of biogenic volatile organic compounds). For my third paper (Gulden et al. 2007a; see Chapter 4), I added a representation of interannual change in leaf mass to the vegetation component of

an LSM. To quantify the importance of interannual differences in biomass to the emission of biogenic volatile organic compounds, I compared the simulations from the original LSM to those of the augmented LSM. To ensure that the dynamic leaf model within the LSM accurately represented interannual leaf variation in Texas, I needed to calibrate many of the model's physical parameters (e.g., minimum leaf size, fraction of assimilated carbon allocated to leaf growth, etc.). Manually calibrating the model brought my attention to the extent to which model output depends on the values of parameters, most of which either cannot be reliably measured or do not correspond to a known physical quantity. Parameter uncertainty is especially pronounced when observational data are scarce and there is consequently not enough information to constrain the model parameters (a situation that is much more common than not).

As I became more aware of the full extent of unavoidable uncertainty in LSM simulations, not just in those of biogenic emissions, I also recognized that LSMs are often developed and evaluated at idealized, data-rich sites but are then applied in regions where validation data are either sparse or unavailable. In my fourth paper, I evaluated the performance of three different representations of subsurface hydrology. I wanted to know whether increasing conceptual realism in a model parameterization decreases a model's sensitivity to errant parameter choices (which are an unavoidable hazard of modeling most locations on Earth because evaluation data are sparse). I quantified the confidence that can be placed in large-scale simulations of change in terrestrial water storage made by LSMs under different levels of parameter uncertainty (Gulden et al., 2007b; see Chapter 5). I also assessed how that confidence varies given changes in the conceptual realism of the LSM.

In an effort to improve both our ability to quantify the uncertainty inherent in LSM simulations and our ability to evaluate models and their component parameterizations, I have proposed a new framework in which to evaluate and develop LSMs, especially when the LSMs are to be used for predictive applications. The framework explicitly takes uncertainty in model parameters, model input data, and evaluation data into account and gives modelers an objective way to quantify how well their model will perform in non-idealized situations (Gulden et al., 2008b; see Chapter 6). The proposed method provides an objective way to quantify the confidence that can be placed in the simulations of LSMs. The framework allows modelers to draw conclusions about model physical parameterizations and model function that do not implicitly depend on their use of ideal parameters, error-free input data, or error-free evaluation data.

Chapter 2: Development of species-based, regional emission capacities for simulation of biogenic volatile organic compound emissions by land-surface models: *A case study of Texas, USA*[1]

2.1. ABSTRACT

A method is introducted to incorporate species-based variation of the emission of biogenic volatile organic compounds (BVOCs) into regional climate and weather models. We convert a species-based land-cover database for Texas into a database compatible with the Community Land Model (CLM) and a database compatible with the Noah land-surface model (LSM). We link the LSM-compatible land-cover databases to the original species-based dataset as a way to derive region-specific BVOC emission capacities for each plant functional type (in the CLM database) and for each land cover type (in the Noah database).

The distribution in space of inherent BVOC flux (defined as the product of the BVOC emission capacity and the leaf biomass density) derived using the Texas-specific BVOC emission capacities correlates well with the spatial distribution of inherent BVOC flux calculated using the original species data (r = 0.89). The mean absolute error for the emission-capacity–derived inherent flux distribution is an order of magnitude lower than the state-wide range of inherent fluxes.

The ground-referenced land-cover datasets derived here are likely more accurate

[1]Substantial portions of this chapter were previously published in Gulden, L.E., and Z.-L. Yang (2006), Development of species-based, regional emission capacities for simulation of biogenic volatile organic compound emissions in land-surface models, *Atm. Env.*, 40, 1464-1479. The References section contains full citations for all articles referenced here.

than their satellite-derived counterparts; they can be used for a variety of regional model simulations in Texas. The inherent BVOC flux distributions derived using region-specific BVOC emission capacities are more consistent with observations than is the BVOC flux distribution derived using the CLM3-standard BVOC emission capacities, which are top-down estimates based on the literature. When used in conjunction with detailed land-cover datasets, region-specific BVOC emission capacities produce reasonably accurate inherent BVOC fluxes.

2.2. INTRODUCTION

Comprehensive, stand-alone models of the climate system require the accurate simulation of processes that involve biogenic emissions. Biogenic volatile organic compounds (BVOCs) serve as messengers between the land surface and atmosphere. Atmospheric conditions influence the rate at which BVOCs are emitted to the atmosphere, and BVOCs, in turn, influence atmospheric chemistry, cloud formation, Earth's radiative balance, and the global carbon cycle.

The long-term pattern of atmospheric variation controls the composition of plant species that covers the land surface; vegetation species exerts primary control on the magnitude of BVOC flux. On a shorter time scale, variation in atmospheric conditions determines the amount of photosynthetic active radiation (PAR) reaching the leaf surface, leaf-surface temperature, and soil moisture, all of which influence variation in BVOC emissions (e.g, Guenther et al., 1991; Plaza et al., 2005). Once emitted to the atmosphere, BVOCs react in the presence of nitrogen oxides to increase the concentration of tropospheric ozone (e.g., Wiedinmyer et al., 2001a; 2001b), which is a respiratory irritant

and major component of smog. The oxidation products of isoprene and non-isoprene BVOCs condense to form secondary organic aerosols (Claeys et al., 2004; Kavouras et al., 1998), which alter Earth's radiative balance and serve as cloud-condensation nuclei (Andreaevc and Crutzen, 1997). Most BVOCs are precursors to atmospheric carbon dioxide and are thereby a way that the land surface and atmosphere exchange carbon (e.g., Guenther, 2002). The magnitude of the BVOC flux between the land-surface and atmosphere is considerable: Guenther and colleagues (1995) estimate that the combined global BVOC emissions contribute 1150 Tg of carbon to the atmosphere each year.

Large-scale study of land-surface–atmosphere feedbacks mediated by biogenic emissions requires BVOC fluxes to be accurately represented within land-surface models (LSMs), which serve as the lower boundary for weather and climate models. The ability of LSMs to correctly represent BVOC flux limits the value of future improvements to weather and climate models' parameterizations of BVOC-related atmospheric processes.

The dependence of BVOC emission on plant species poses a significant obstacle to the accurate representation of biogenic emissions within LSMs. Although changes in environmental conditions cause the greatest variation in the quantity of BVOCs emitted from a single plant, interspecies variation in BVOC emission capacity exerts primary control on the differences in the rates at which disparate landscapes emit BVOCs. Current-generation LSMs represent vegetation either as static land-cover classes (e.g., as in the Noah LSM [Chen and Dudhia, 2001]) or as mosaics of plant functional types (PFTs), such as "temperate needleleaf evergreen tree" and "temperate broadleaf deciduous shrub" (e.g., as in the Community Land Model [CLM] version 3 [Oleson et al., 2004; Bonan et al., 2002]). The vast number of species represented by a single LSM land-

cover type makes accurate representation of BVOC fluxes using LSM land-cover types a challenge.

In their BVOC emissions module, which is standard in CLM3, Levis et al. (2003) (hereafter, "Levis et al.") represent broad-scale interspecies variability in BVOC emissions by assigning unique emission capacities to each PFT. They base their emissions module on the model of Guenther et al. (1995), who represented BVOC flux as a function of foliar density, PAR, leaf-surface temperature, and an ecosystem-specific emission capacity. For a given PFT, Levis et al. represent the emission of BVOC type *i* (isoprene, monoterpene, other volatile organic compounds [OVOCs], other reactive volatile organic compounds [ORVOCs], or CO) as:

$$F_i = \varepsilon_i D \ (L_{sun} + L_{shade})_i T_i \qquad (2.1)$$

where F_i is the flux to the atmosphere of BVOC type *i* (units: μg C m^{-2} h^{-1}); D is the foliar density (units: g dry leaf matter [gdlm] m^{-2} of ground covered by the PFT); ε_i is a PFT-specific emission capacity (units: μg C $gdlm^{-1}$ h^{-1}); L_{sun} and L_{shade} are dimensionless factors that modulate BVOC emissions from sunlit and shaded leaves in response to changes in PAR; and T_i is a dimensionless factor that adjusts BVOC flux in response to changes in canopy temperature.

Globally observed BVOC emissions data do not yet exist; however, the magnitude and spatial variability of the global BVOC flux simulated offline by Levis et al.'s module are consistent with other model estimates (e.g., Guenther et al., 1995). When Levis et al. simulated BVOC emissions at 3.75º×3.75º resolution for ten model years using the fully coupled Community Climate System Model (CCSM) with dynamic

vegetation, interannual variability in total BVOC flux reached 29%. This is consistent with the work of Abbot and colleagues (2003), who used space-based atmospheric column measurements of formaldehyde as a proxy for isoprene emissions over North America. The scientists found that between 1995 and 2001, interannual variability in estimated isoprene emissions was approximately 30% (Abbot et al., 2003); the spatial distribution and magnitude of their estimated isoprene emissions generally matched the predictions of models with emissions parameterizations similar to that of Levis et al. Plot-scale observational studies also report appreciable interannual variation in BVOC emission (e.g., Hakola et al., 2003).

Levis et al. used PFT-specific emission capacities that were globally constant. The top-down assignment of PFT-specific emissions capacities is a simplification that is defensible in the case of qualitative global simulations but is likely insufficient for regional simulations. Levis et al. suggest using region-specific emission capacities as a way to increase the accuracy with which CLM3 simulates BVOC fluxes on a regional scale. Here we present a "bottom up" method for determining Texas-specific BVOC emission capacities for PFTs using a species-based, ground-referenced land-cover database (Wiedinmyer et al., 2001b). We also apply this method to develop regional BVOC emission capacities for the static land-cover types used in the Noah LSM, which is used as the lower boundary for the fifth-generation Penn State/NCAR Mesoscale Model (MM5) (Grell et al., 1995) and for the Weather Research and Forecasting Model (WRF) (Skamarock et al., 2005). CLM is an LSM representative of those used in climate model simulations; the Noah LSM is widely used in both the weather- and climate-

modeling communities. Methods described here can be applied to other LSMs as research needs dictate.

Within the work presented here, we address only the BVOC emission capacity of the land surface. We do not report simulation results. This study makes the following unique contributions: (1) the development of two ground-referenced, species-based, 1-km land-cover databases for Texas that can be used as input to LSMs (CLM and Noah) and (2) the demonstration and evaluation of a method for deriving region-specific BVOC emission capacities for use in regional climate and weather models.

2.3. METHODS AND DATA

2.3.1. The Wiedinmyer database

The Wiedinmyer dataset describes the Texas landscape using 600+ species-based land-cover classes (Wiedinmyer et al., 2001b). Wiedinmyer and colleagues developed the database as input to an MS Access–based BVOC emissions module, GLOBEIS (http://www.globeis.com; Yarwood et al., 1999). In places where BVOC emissions are typically high (in eastern and central Texas), the researchers conducted field surveys to determine species composition and density information. The researchers then mapped that information to the land cover classes contained in ten existing, often overlapping land-cover databases for various regions in Texas. A detailed description of the methods used to develop the dataset can be found in Wiedinmyer et al. (2000; 2001b).

The resolution of the Wiedinmyer database is approximately 1 km; some urban areas have a finer resolution. We employed ESRI's ArcGIS software to convert the original Wiedinmyer database to a uniform 1×1-km gridded dataset.

Each land-cover class in the Wiedinmyer dataset contains species composition information. One class contains between 1 and 115 of 289 possible component species. For every species that makes up a land-cover class, the database contains the percent of the land-cover class ground area that is covered by the species and the foliar density of the species, D_i, which is the leaf biomass of the given species that is contained in each unit area of the land cover class. D_i is defined as:

$$D_i = \frac{A_i d_i}{A_{total}} \tag{2.2}$$

where d_i is the leaf biomass of species i per unit area covered by species i, A_i is the land-cover–class ground area covered by species i, and A_{total} is the total ground area covered by the land-cover class. ($A_i\, A_{total}^{-1}$ is the percent of the area of the given land-cover class that is covered by species i.) The sum of the area fraction of all species in a land-cover class is the total percent vegetated area; the remaining percentage is assumed unvegetated. When linked to tables contained in GLOBEIS, the Wiedinmyer database also contains BVOC emission capacities (units: μg C gdlm^{-1} h^{-1}) for each of the contained 289 vegetation species.

2.3.2. Conversion of Wiedinmyer database to plant functional types

We used the Wiedinmyer database to develop a dataset that describes the Texas landscape in terms of CLM-compatible PFTs (Figure 2.1). Every species contained in the Wiedinmyer database was assigned to one PFT. We assumed all PFTs in Texas are temperate (i.e., because the state is entirely located in the midlatitudes, there are neither

tropical PFTs nor boreal PFTs in Texas). A list of definitions of PFTs present in Texas can be found in the footnotes of Table 2.1.

Species labeled as part of phylogenetic divisions Cycadophyta, Ginkgophyta, and Magnoliophyta were considered broadleaf; species in division Coniferophyta, needleleaf. To determine whether a species is evergreen or deciduous, we used the USDA Plants National Database (http://plants.usda.gov) as our primary source of information. Secondary sources included the Texas A&M Trees and Shrubs of Texas databases (http://aggie-horticulture.tamu.edu/ornamentals/natives/tamuhort.html) and the work of Stahl and McElvaney (2003). If information about the leaf life of a species was contradictory, priority was given to the information in the Plants National Database. If information was unavailable, the species was considered deciduous or evergreen based on the characteristics of plants sharing the same genus and based on local knowledge.

The database developed by Wiedinmyer does not contain information that allows for an objective description of species as shrubs or trees. In the initial conversion, all shrubs were designated trees. For example, *Atriplex canescens*, four-wing saltbush, although morphologically a shrub, was considered a broadleaf deciduous tree (BDT) in the first analysis. After the initial conversion, all "trees" in Wiedinmyer land-cover classes whose descriptive names contained the words "Shrub," "Brush," or "Shrub and Brush" were reclassified as shrubs. (One should note that this method prevents the coexistence of trees and shrubs in a given 1×1-km grid cell, which likely misrepresents reality in transitional regions such as the savannas of central Texas.) The standard initialization of the CLM does not contain a temperate needleleaf evergreen shrub (NES), but the prevalence of temperate evergreen coniferous shrub species (e.g., Alligator

Juniper [*Juniperus deppeana*]) in some regions of western Texas justified the inclusion of NES as a component PFT. The standard initialization of CLM does not contain a temperate needleleaf deciduous tree (NDT), but the presence of bald cypress (*Taxodium distichum*) in eastern Texas justified the addition of temperate NDT to the PFTs existing in Texas.

Regardless of photosynthetic pathway, all crop species were considered part of the same PFT (crop). Because the Wiedinmyer database does not provide species information for vegetation identified as "grass," we were unable to use the Wiedinmyer database to distinguish between C_3 and C_4 grasses. To define grass as C_3, C_4, or a C_3–C_4 mix, we followed the methods of Bonan et al. (2002). Grass populations in locations where the mean monthly temperature never reaches or exceeds 22ºC were defined as pure C_3. No location in Texas meets the criteria to contain pure C_4 grass, which requires that mean monthly temperatures exceed 22ºC throughout the year. Remaining grasslands were considered to be 50% C_3 and 50% C_4. To approximate monthly mean temperature, we averaged PRISM mean monthly maximum and minimum temperature data (Spatial Climate Analysis Service, 2004). The original resolution of the PRISM grids was 4 km; we used nearest-neighborhood interpolation to obtain 1-km grids.

For every land-cover class in the Wiedinmyer datasets, we calculated (1) the percent of ground area covered by every PFT and (2) the foliar density of each component PFT. The former was calculated by summing the area fraction (units: $m^2\ m^{-2}$) of all species in the class identified as a PFT. The later was defined as the sum of the foliar densities of all plant species contained within the land-cover class identified as a PFT. Figure 2.1 shows a schematic diagram of the process, and Figure 2.2 displays the

resulting PFT distribution. We employed information contained within the Wiedinmyer database to created gridded landunit-level maps of urban and water-covered areas in Texas.

2.3.3. Conversion of Wiedinmyer database to land-cover types for the Noah LSM

Similar to many LSMs, Noah uses static land-cover types (e.g., "Mixed Forest," "Cropland/Grassland Mosaic") to represent surface vegetation. Because a Noah land-cover type can contain more than one type of plant (e.g., "Savanna" contains both grass and tree species), conversion from the species-based Wiedinmyer dataset directly to Noah land-cover types could be accomplished only by using subjective methods. We instead determined the distribution of Noah land-cover types in Texas by using the PFT percent composition information as input to a decision-tree algorithm (Figure 2.3). The resulting database can be used as input to WRF–Noah version 2.1, the default version of which uses land-cover categories identical to those used in the USGS Land Use/Land Cover System (Anderson et al., 1976). Figure 2.5 compares the Wiedinmyer-derived Noah land cover with the default USGS-based land cover.

Breakpoint values within the decision tree (e.g., "MinValTrees," the minimum percentage of tree cover required for a land-cover class to be considered a type of forest, savanna, wooded wetland, or a cropland–woodland mosaic; see legend of Figure 2.3) were chosen so that the definitions of the land-cover classes were broadly consistent with the work of other researchers (e.g., Hansen et al., 2000).

2.3.4. Calculation of Texas-specific BVOC emission capacities

The emissions module within CLM calculates BVOC fluxes for isoprene, monoterpene, and OVOCs, as well as for CO and ORVOCs. In the standard version of CLM3, emission capacities for CO and ORVOCs do not differ between PFTs. The GLOBEIS-linked Wiedinmyer database contains plant-species–specific emissions rates for only isoprene, monoterpene, and OVOCs. We calculated region-specific emission capacities for only these three BVOC species. The emission capacity of PFT_x with respect to a single BVOC type, ε_{PFTx} (units: μg C gdlm^{-1} h^{-1}), was calculated as:

$$\varepsilon_{PFT_x} = \frac{F_{PFT_x}}{W_{PFT_x}} \tag{2.3}$$

where F_{PFT_x} is the total BVOC emissions per hour emitted by PFT_x across Texas (units: μg C h^{-1}) and W_{PFTx} is the total leaf biomass (gdlm) of PFT_i in Texas. We defined F_{PFT_x} as:

$$F_{PFT_x} = \sum_{j=1}^{n} A_j \left(\sum_{i=1}^{m_{PFT_x}} \varepsilon_i D_i \right)_j \tag{2.4}$$

where n is number of land-cover classes in Texas; A_j is the total area in Texas covered by land-cover class j; m_{PFT_x} is the number of species in land-cover class j identified as PFT_x; ε_i is the emissions rate per unit biomass (units: μg C gdlm^{-1} h^{-1}) of the ith species identified as PFT_x in land-use code j; and D_i is the foliar density (as defined in Equation 2.2) of the ith species in land-use code j. Figure 2.4 shows the distribution of leaf biomass density, according to the Wiedinmyer database. We calculated W_{PFTx} as:

$$W_{PFT_x} = \sum_{j=1}^{n} A_j \left(\sum_{i=1}^{m_{PFT_x}} D_i \right)_j \qquad (2.5)$$

where A_j, D_i, and m_{PFT_x} are as described above.

Significant variation exists within the emission capacity of the Texas-native plant species that compose each PFT. It is possible that plant species' emission capacities vary regularly as a function of location within Texas. (For instance, species classified as BDT that are native to western Texas may have emission characteristics that differ significantly from those of species classified as BDT that are native to eastern Texas.) To explore this possibility, we derived two sets of region-specific PFT emission capacities for two sub-regions in Texas (eastern and western Texas). 99ºW approximately bounded the two sub-regions.

Table 2.1 compares the emission capacities of Levis et al. with the region-specific emission capacities derived here. Following methods similar to those outlined above, we also derived region-specific emission capacities for use with the Noah land-cover dataset (Table 2.2).

2.3.5. Computation of specific leaf area for use in applications of LSMs

The Levis et al. emissions module calculates BVOC emissions as a function of leaf biomass density, but CLM tracks variations in vegetation biomass using leaf area index (LAI) (units: m^2 leaf m^{-2} ground). Levis et al. assign each PFT a time-invariant specific leaf area (units: m^2 leaf $gdlm^{-1}$), which allows for conversion between LAI and leaf biomass density (Foley et al., 1996; Levis et al., 2003). Using static LAI and leaf biomass density values provided in the Wiedinmyer database, we derived specific leaf

area (SLA) for each PFT and each Noah land-cover type using a biomass-weighted averaging method analogous to that used to derive region-specific BVOC emission capacities (Table 2.3).

SLA values are not necessary for evaluation of the region-specific emission capacities; however, we present them here because they are necessary components of model simulations that use the Texas-specific emission capacities derived in this study. It should be noted that the Noah LSM uses greenness fraction, not LAI, to track variations in biomass. The SLA value for each Noah land-cover code shown in Table 2.3 must therefore be used with a mechanism that converts greenness fraction to LAI.

2.3.6. Calculation of inherent BVOC flux

The inherent BVOC flux was used as a measure by which to evaluate the accuracy of the emission capacities derived here. Inherent flux, *F,* for each type of BVOC was defined as:

$$F = \sum_{i=1}^{n} \varepsilon_i D_i \qquad (2.6)$$

Where n is the number of species or land-cover types covering the grid cell; ε_i is the emission capacity of species, land-cover type, or PFT i; and D_i is the foliar density of i (as defined in Equation 2.2). In all cases, we used the foliar density information provided in the Wiedinmyer database to compute inherent BVOC flux. For computations using PFTs, we considered D_i to be the sum of the foliar densities of all species in a grid cell identified as PFT i. Because Noah land cover classes are a static composition of plant types (n = 1 in Equation 2.6), the inherent flux for a given 1×1-km grid cell is $F = \varepsilon D$,

where ε is the emission capacity for the land-cover class, and D is the grid-cell–mean foliar density (as defined in Equation 2.2). Evaluation of results using inherent BVOC flux calculated using a standard leaf biomass density map allowed us to focus entirely on variation in BVOC emission resulting from differences in emission capacities, neglecting differences resulting from environmental variation.

2.4. RESULTS AND DISCUSSION

The total inherent BVOC flux was defined as the sum of the inherent fluxes for isoprene, monoterpene, and OVOCs. Figure 2.6 shows the state-wide total inherent BVOC fluxes calculated using the Levis et al. emission capacities and the Wiedinmyer-derived CLM-compatible land-cover dataset ("Levis"); calculated using the one-region, Texas-specific PFT emission capacities and the CLM-compatible dataset ("One-region"); calculated using the two-region, Texas-specific PFT emission capacities and the CLM-compatible dataset ("Two-region"); and calculated using the Texas-specific Noah land-cover emission capacities and the Noah-compatible land-cover dataset ("Noah"). When evaluating the ability of the region-specific emission capacities to represent BVOC emissions in Texas, the inherent BVOC flux derived directly from the species-specific emission capacities and the species-based land-cover map were considered to be the "observed" data ("Wiedinmyer" in Figure 2.6). Although the distribution in space of the "Levis" inherent fluxes is qualitatively similar to that of the observed data, the inherent fluxes calculated with region-specific emission capacities greatly improve the LSMs' ability to simulate BVOC emissions.

Table 2.4 shows univariate statistics for the results shown in Figure 2.6; Table 2.5 displays the correlation coefficient, the mean absolute error (MAE), and the root mean square error for the results shown in Figure 2.6. The inherent BVOC fluxes calculated using the region-specific emission capacities are better correlated with the observed fluxes than are the inherent fluxes derived using Levis et al. emission capacities. A comparison of the MAE values for the inherent flux datasets shows that regional emission capacities greatly improve the ability of LSMs to simulate inherent BVOC fluxes, and consequently, BVOC emissions.

All MAE values for the region-specific inherent fluxes are 41% or less of the mean inherent flux, and all are an order of magnitude lower than the range of values in the observed datasets. This compares quite favorably to the MAE of the "Levis" inherent emission rates, which is 73% of the dataset's mean value and 19% of the range. When the Texas-specific emission capacities derived in this study are used in climate and weather simulations, the spatial distribution of BVOC emissions will better reflect observed emissions, which will consequently improve the spatial accuracy of the modeled climate and weather processes that depend on biogenic emissions (e.g., secondary organic aerosol formation and tropospheric ozone production).

Globally constant emission capacities from Levis et al. yield substantial underestimates of the total BVOC emissions in Texas: the cumulative state-wide inherent BVOC flux calculated using the Levis et al emission capacities was only 85% of the observed (Wiedinmyer) cumulative state-wide inherent flux. Because the method presented here preserves the total mass of BVOCs emitted from the landscape under normalized temperature and PAR conditions, each of the three Wiedinmyer-derived

inherent flux distributions has cumulative state-wide inherent BVOC fluxes that are identical to that computed using the original Wiedinmyer dataset (Table 2.4). However, the different spatial distribution of inherent fluxes obtained using the Wiedinmyer-derived emission capacities results in different average BVOC fluxes at regional scales. Figure 2.7 shows the average inherent flux of BVOCs for four major metropolitan areas in Texas. In each of the four urban areas, use of any of the three Wiedinmyer-derived emission capacities ("One-region," "Two-region," and "Noah" in Figure 2.7) substantially improves upon the average inherent flux calculated using Levis et al.'s values ("Levis" in Figure 2.7).

Use of emission capacities derived from two sub-regions improves the spatial distribution of simulated inherent emission capacities, especially in central Texas; however, the improvement in accuracy is not as great as one might expect. The lower foliar density of the west-Texas landscape (Figure 2.4) likely accounts for a greater portion of the difference in the inherent fluxes than do the slight variations in the regional emission capacities between the species native to east and west Texas (Table 2.1). Different choice of sub-regions (e.g., north and south Texas) might also provide a greater increase in accuracy.

Further improvement of BVOC-emission simulation could likely be gleaned by the addition of "high-emitting" land-cover types to CLM and Noah. For example, a BDT that shares all physical parameters except its isoprene emission capacity with the CLM-standard BDT would represent BDT species with anomalously high isoprene emission capacities (e.g., oak species). Emission capacities calculated for the "high-isoprene" BDT and the standard BDT would consequently be more representative of the emission

capacities of their constituent species. Similar high-emitting PFTs for plant species with anomalous monoterpene or OVOC emission capacities would improve model simulation results in regions where such species are prevalent. Future work will include the addition of such land-cover types to both CLM and Noah.

We focus on the relative improvement gained by using region-specific, species-based emission capacities. Evaluation of the absolute improvement in accuracy gained by using the ecologically based emission capacities is challenging because of a dearth of large-scale, high-resolution empirical data. However, the ecologically based, "bottom up" method presented here is presumably more plausible than the "top down" assignation of emission capacities to land-cover types used in similar modeling studies (e.g., Levis et al., Naik et al. , 2004). Means for evaluation of the absolute accuracy of the magnitude and spatial distribution of BVOC emissions and the weather and climatic processes that depend upon them is an area of ongoing research.

Application of this method to other regions may be hindered by a dearth of detailed, species-based land-cover databases. Less-detailed species-based vegetation datasets, such as the 1-km Biogenic Emissions Landcover Database (Kinnee et al., 1997), can be used in similar bottom-up fashion to derive region-specific emission capacities for LSMs. In the United States, myriad less extensive vegetation species databases sponsored by state and federal agencies (e.g., the Calflora vegetation species database [http://www.calflora.org/]) augment the species-distribution data available to researchers.

The BVOC emission capacities derived for Texas may be representative of those in humid-to-arid transition zones and may thus be transferable to regions with similar climate. Sen and colleagues (2001) used observational data from five sites, each

representative of a different biome, to calibrate the vegetation parameters for a general circulation model; their model simulations using the field-calibrated vegetation data showed more skill than did those using standard parameters. A similar approach using BVOC emission capacities derived from representative regions might improve modeled BVOC emissions in continental- and global-scale simulations without requiring detailed, species-based vegetation maps for all regions. However, because BVOC emissions vary at the species level, extensive further study is necessary to determine whether such generalization is defensible.

As a derivative product, this study inherits all uncertainty associated with the input Wiedinmyer land-cover dataset. As Wiedinmyer and colleagues assert (2000, 2001b), the uncertainty associated with the original dataset is difficult to quantify. The extensive field surveys undertaken by Wiedinmyer and colleagues and the field-surveys underpinning the ten land-cover datasets that formed the basis for the Wiedinmyer dataset serve to lessen the associated uncertainty; however, field-survey verification of the Wiedinmyer dataset was limited to central and eastern Texas, where BVOC emissions rates are high. Although ground surveys for the datasets underpinning the Wiedinmyer dataset were generally not as localized, the uncertainty of the Wiedinmyer dataset increases from east to west (Wiedinmyer et al., 2001b).

The accuracy of the inherent BVOC fluxes calculated here depends on the accuracy of the emissions data contained within GLOBEIS. Intraspecies variation in BVOC emissions and the dependence of emissions on environmental variables lessens the transferability of site-specific, BVOC-emissions measurements to regional modeling applications. Intraspecies variation is likely considerable: for example, Funk and

colleagues (2005) found that under normal climate conditions, isoprene emissions from different red oak trees within a single stand differed by a factor of two. Interspecies variation also introduces uncertainty. The number of vegetation species in Texas is considerably greater than the 289 species contained within the Wiedinmyer database; for most species in Texas, observed BVOC emission capacities do not exist. New field and laboratory data (e.g., Geron et al., 2001) combined with methods for assigning BVOC emission capacities to species for which no measurements exist (e.g., Karlik et al., 2002) will lessen but not eliminate this uncertainty.

Another source of uncertainty inherent to the method presented here results from the use of a time-varying property (leaf biomass density) to derive a time-invariant property (PFT-specific emission capacity). The emission capacities for use with the Noah-compatible database also depend on the selection of the breakpoint values in the decision-tree algorithm used to derive the Noah-compatible database from the PFT database. Detailed factorial analysis could help determine the most accurate breakpoint values (e.g., Henderson-Sellers, 1993).

2.5. SUMMARY AND CONCLUSIONS

The CLM-compatible and Noah-compatible datasets developed for this study can be used for a wide variety of regional weather and climate simulations. They are the first high-resolution, ground-referenced vegetation datasets for Texas for use in land-surface models.

The use of CLM3-standard PFT emission capacities likely yields underestimate of the cumulative BVOC emissions from vegetation in Texas and fails to capture the range

of inherent BVOC fluxes. Region-specific emission capacities, calculated using either one region or two smaller sub-regions, improve the ability of LSMs to simulate the inherent BVOC flux. The correlation between the observed inherent fluxes and the inherent fluxes calculated using region-specific emission capacities (r = 0.89) is better than that between observations and the fluxes calculated using the CLM standard emission capacities (r = 0.85). Use of region-specific emission capacities lessens the mean absolute error of the calculated inherent flux. (Mean absolute error is 1668 μg C m^{-2} h^{-1} when using standard CLM emission capacities; 1111 and 1085 μg C m^{-2} h^{-1} when using region-specific CLM emission capacities derived from one and two regions, respectively; and 945 μg C m^{-2} h^{-1} when using region-specific Noah emission capacities.) Subdivision of Texas into two smaller sub-regions improved the accuracy of the spatial distribution of inherent emission capacities. However, for regions the size of Texas, a single set of regional PFT BVOC emission capacities may be used without a large compromise in the overall accuracy of the spatial distribution of emissions.

When used in conjunction with detailed land-cover datasets, region-specific, species-derived BVOC emission capacities allow for reasonably accurate simulation of BVOC emissions on a regional scale. Climate and weather models that employ region-specific BVOC emission capacities will be better equipped to study the impact of biogenic emissions on atmospheric processes and the influence of atmospheric variation on biogenic emissions.

2.6. ACKNOWLEDGEMENTS

The Environmental Protection Agency provided funding for this research (grant no. RD83145201). We thank Christine Wiedinmyer, Victoria Junquera, and William Vizuete for providing us with the original dataset and for their insight and assistance over the course of the research. We would also like to thank Robert Dickinson for valuable comments regarding methods for improving simulation of BVOC emissions in LSMs. This manuscript was substantially improved by the insightful comments of two anonymous reviewers and of Lauren Greene.

Table 2.1. Comparison of BVOC emission capacities (ε_i, units: μg C gdlm^{-1} h^{-1}) and specific leaf area (SLA) values for PFTs.*

PFT → / Method ↓	NET	NES†	NDT†	BET	BES	BDT	BDS	Crop	C3,C4 Grass	Water‡
	Isoprene									
Levis et al.	2	2	0	24	24	24	24	0	0	0
CLM3: 1 region	0.1	0.1	0.2	44.6	12.6	43.1	3.8	0	0	2.5
CLM3: 2 regions West	0.1	0.1	0	44.8	13.5	26.2	4.1	0	0	0
CLM3: 2 regions East	0.1	0.1	0.2	44.6	0.8	45.5	2.2	0	0	2.5
	Monoterpene									
Levis et al.	2	2	1.6	0.8	0.8	0.8	0.8	0.1	0.1	0
CLM3: 1 region	2.2	0.5	2.3	0.5	0.7	0.6	0.6	0.1	0	1.6
CLM3: 2 regions West	1	0.5	0	1.5	0.7	0.8	0.7	0.1	0	0
CLM3: 2 regions East	2.4	0.4	2.3	0.3	0	0.6	0	0.1	0	1.6
	OVOC									
Levis et al.	1	1	1	1	1	1	1	1	1	0
CLM: 1 region	1.9	1.3	1.3	2	0.7	2.1	1.1	0	0	1.8
CLM3: 2 regions West	2.1	1.3	0	2	0.7	2.8	1.3	0	0	0
CLM3: 2 regions East	1.9	1.1	1.3	2	0.1	2	0.1	0.1	0	1.8

* Emission capacities shown are rounded to the nearest tenth, but non-rounded emission capacities were used. NET is needleleaf evergreen tree; NES is needleleaf evergreen shrub; NDT is needleleaf deciduous tree; BET is broadleaf evergreen tree; BES is broadleaf evergreen shrub; BDT is broadleaf deciduous tree; BDS is broadleaf deciduous shrub.

† For the calculations of state-wide inherent fluxes using Levis et al.'s emission capacities, the NES emission capacity was assumed to have the same value as the Levis et al. emission capacity for NET. NDT (a temperate PFT) was assumed to have the same value as the Levis et al. emission capacity for boreal NDT.

‡ Water is not a PFT in CLM; however, in the Wiedinmyer database, water bodies can emit BVOCs. Consequently, an emission capacity for water was calculated.

Table 2.2. BVOC emission capacities (ε_i, units: μg C gdlm^{-1} h^{-1}) for Noah land-cover types.

Land cover*	Isoprene†	Monoterpene†	OVOC†
Urban	30.1	0.8	1.7
Crop	0.1	0.1	0.1
Crop/Grass	0	0.1	0
Crop/Wood	12.1	0.3	0.7
Grass	0	0	0
Shrub	8	0.9	1.4
Shrub/Grass	0.1	0.2	1.3
Savanna	0	0	0
DBF	36.7	0.5	1.8
ENF	12.2	2.6	1.8
Mixed	15	0.6	1.8
Water	2.2	1.4	1.6
Herbaceous Wetland	1.4	0.9	1.1

* Urban is Urban/Built up; Crop is Irrigated/dryland cropland; Crop/Grass is Cropland/grassland mosaic; Crop/Wood is Cropland/woodland mosaic; Shrub is Shrubland; Shrub/Grass is Shrubland/grassland mosaic; DBF is Deciduous broadleaf forest; ENF is Evergreen needleleaf forest. DNF (Deciduous needleleaf forest) and EBF (Evergreen broadleaf forest) are not included in the above table because, in this analysis, nowhere in Texas is identified as EBF or DNF. "Mixed" is Mixed forest, which is any mixture of DBF, EBF, ENF, and/or DNF, in which no tree type covers a percentage of the vegetated area that exceeds MinValPure (see Figure 2.3).

† Emission capacities shown are rounded to the nearest tenth, but raw values for emission capacities were used to calculate state-wide inherent emission rates.

Table 2.3. Specific leaf area (SLA) (m^2 leaf $gdlm^{-1}$) for CLM plant functional types and Noah LSM land-cover types.

		CLM plant functional type								
		NET	NES†	NDT†	BET	BES	BDT	BDS	Crop	C3,C4 Grass
Method	Levis et al.	0.013	0.013	0.013	0.025	0.025	0.025	0.025	0.02	0.02
	CLM3: 1 region	0.014	0.014	0.024	0.032	0.019	0.03	0.016	0.031	0.02‡
	CLM3: 2 regions West	0.013	0.012	0.024	0.029	9e-4	0.028	0.001	0.029	0.02‡
	CLM3: 2 regions East	0.022	0.014	n/a	0.048	0.02	0.042	0.018	0.031	0.02‡

	Noah LSM land-cover type*										
Urban	Crop	Crop/ Grass	Crop/ Wood	Grass§	Shrub	Shrub/ Grass	Savan -na	DBF	ENF	Mixed	Herb Wet.
0.023	0.03	0.022	0.028	0.006	0.023	0.019	0.024	0.026	0.009	0.022	0.039

***** See first footnote of Table 2.1 for full names corresponding to the PFT acronyms; see first footnote of Table 2.2 for full names corresponding to land-cover type abbreviated names.

† Temperate NDT and NES do not exist in the standard version of CLM3; specific leaf area for NDT and NES are assumed to be equal to the Levis et al. specific leaf area for NET.

‡ The specific leaf area calculated for grass species using the data provided in the Wiedinmyer database (SLA ≈ 0.00001) was unreasonably small; consequently, we recommend the use of the standard-CLM specific leaf area for grass.

§ Because the specific leaf area calculated for grass is unreasonably small (see previous footnote), the specific leaf area calculated for grasslands (which contain a large percentage of grass species) may also be a significant underestimate.

Table 2.4. Univariate statistics for inherent BVOC* fluxes in Texas

Method used to calculate inherent fluxes	Range of inherent BVOC fluxes (μg C m^{-2} h^{-1})	Mean inherent flux (μg C m^{-2} h^{-1})	SD	Skewness	Kurtosis
Wiedinmyer database ("observations") calculated directly from species	0–20,737	2689	3973	1.8005	2.7289
CLM: Levis et al. emission capacities (globally constant)	0–8887	2285	1928	0.892	0.337
CLM: Texas-specific emission capacities (1 region)	0–15,567	2689	3453	1.4303	1.3434
CLM: Texas-specific emission capacities (2 regions)	0–16,156	2689	3531	1.5253	1.6656
Noah: TX-specific emission capacities	0–18,131	2689	3360	0.0165	–0.7756

* Inherent BVOC fluxes are the sum of the inherent fluxes of isoprene, monoterpene, and OVOC.

Table 2.5. Correlation and error of calculated inherent BVOC fluxes.

Method used to calculate BVOC emissions ↓	Pearson's product-moment correlation coefficient, *r*	Mean absolute error*	Root mean square error*
Calculated directly from species-specific emission capacities	1	0	0
CLM: Levis et al. emission capacities (globally constant)	0.8502	1668	2577
CLM: TX-specific emission capacities (1 region)	0.8875	1111	1832
CLM: TX-specific emission capacities (2 regions)	0.8899	1085	1813
Noah: TX-specific emission capacities	0.8893	945	1629

* As defined in Willmott, 1982.

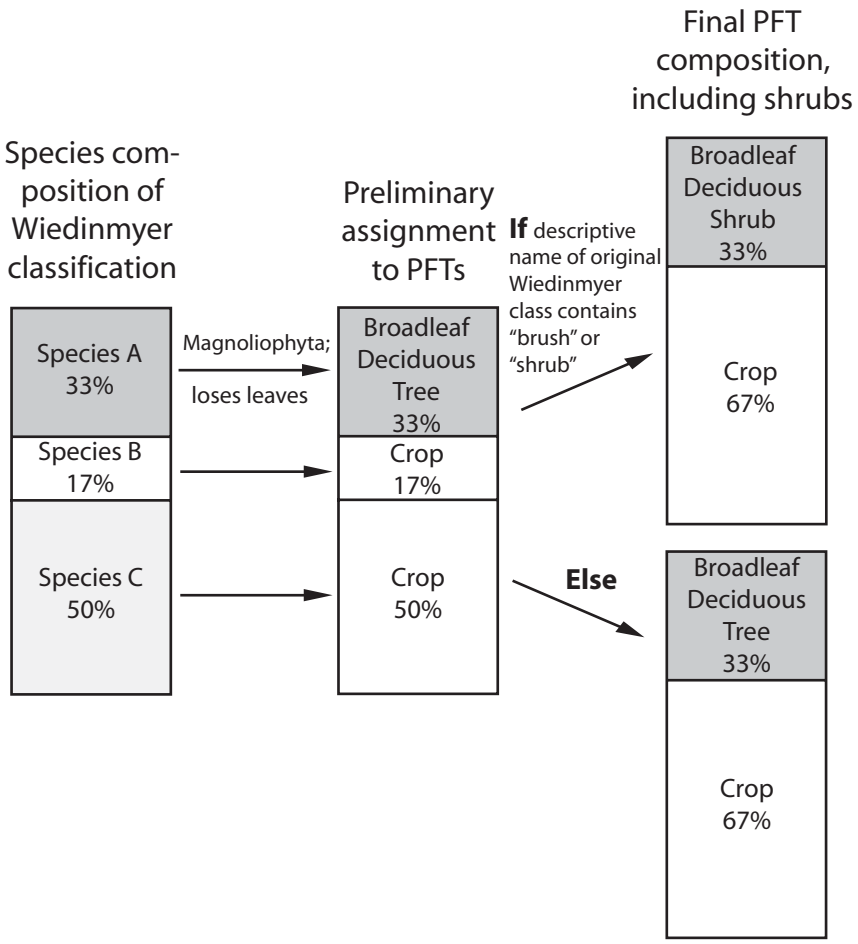

Figure 2.1. Schematic diagram of the PFT conversion process.

Percent area shown is percent of vegetated area. Schematic representation of the calculation of foliar density and total vegetated area is not included.

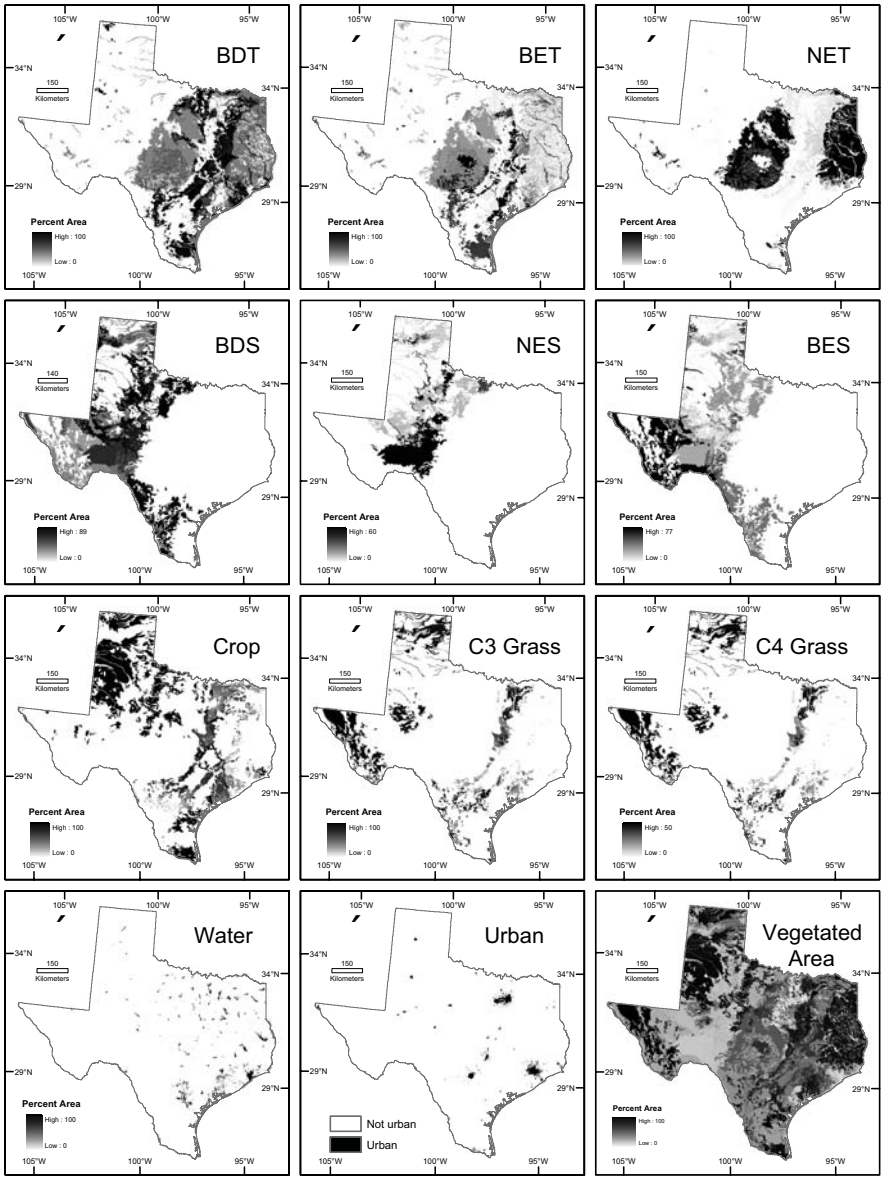

Figure 2.2. Percent of vegetated area covered by different PFTs.

The lower right-hand panel shows percent total vegetated area. NET is needleleaf evergreen tree; NES is needleleaf evergreen shrub; BET is broadleaf evergreen tree; BES is broadleaf evergreen shrub; BDT is broadleaf deciduous tree; BDS is broadleaf deciduous shrub. Because needleleaf deciduous trees are rare in Texas, their distribution is not shown. NES is not a PFT in the standard version of CLM3; see preceding section for discussion. *, ** Water and Urban are landunit-level land-covers in CLM, not PFTs.

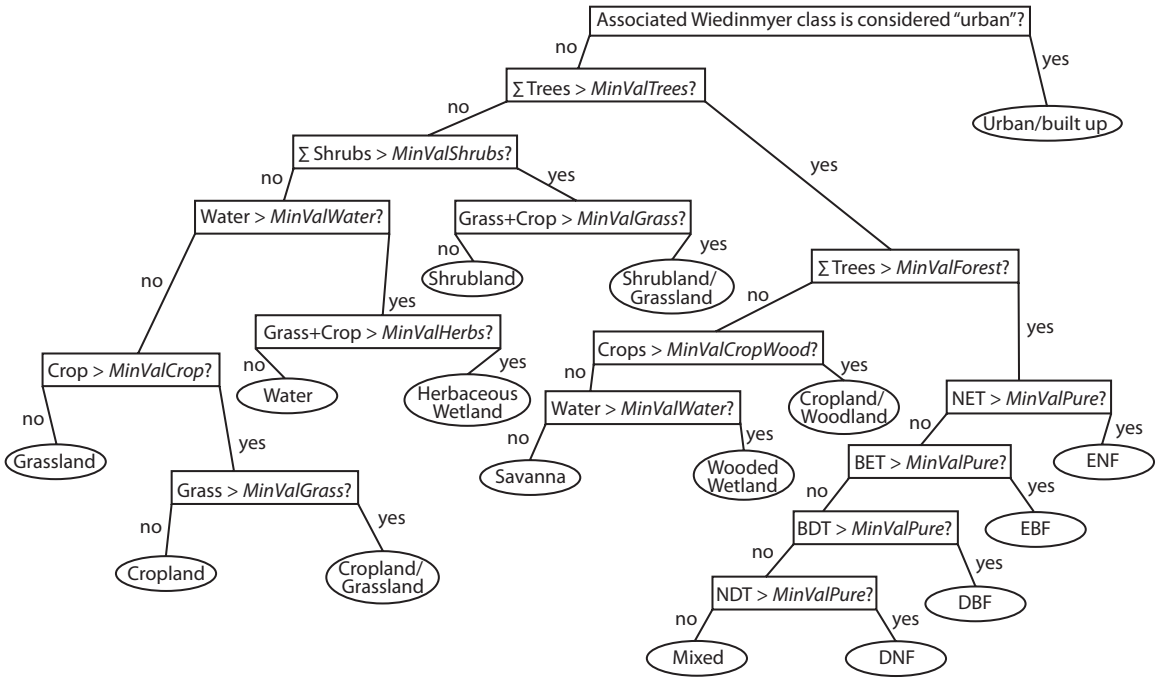

Figure 2.3. Decision tree used to derive a Noah-compatible land-cover classification from the Wiedinmyer-based PFT database.

ENF is evergreen needleleaf forest; EBF is evergreen broadleaf forest; DBF is deciduous broadleaf forest; and DNF is deciduous needleleaf forest. "Mixed" is forest containing a mixture of trees in which no type (ENF, EBF, DNF, or DBF) exceeds the value chosen for "MinValPure." Breakpoint values are as follows: MinValTrees = 10; MinValForest = 50; MinValCropWood = 40; MinValShrubs = 10; MinValWater = 70; MinValCrop = 10; MinValGrass = 20; MinValHerbs = 10; MinValPure = 60

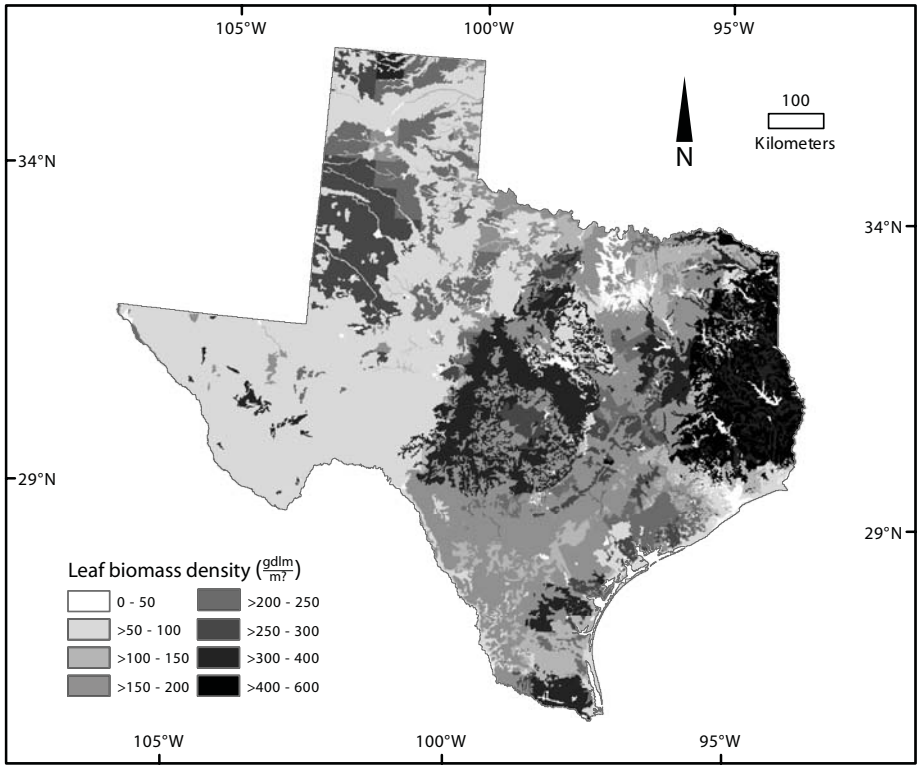

Figure 2.4. Leaf biomass density (gdlm m^{-2}) according to the Wiedinmyer database.

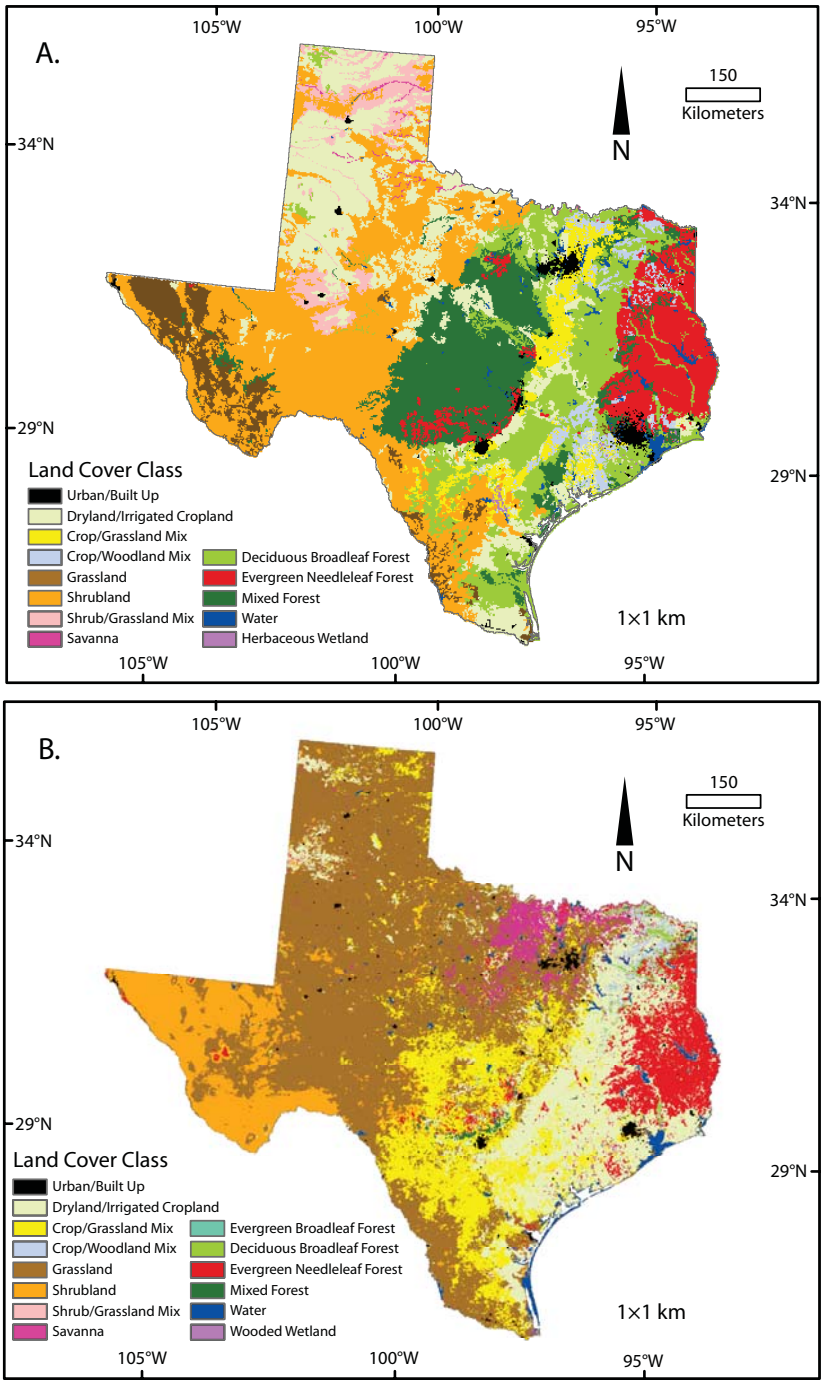

Figure 2.5. Results of conversion of Wiedinmyer database to Noah land-cover (Panel A); default USGS-based Noah land-cover distribution (Panel B).

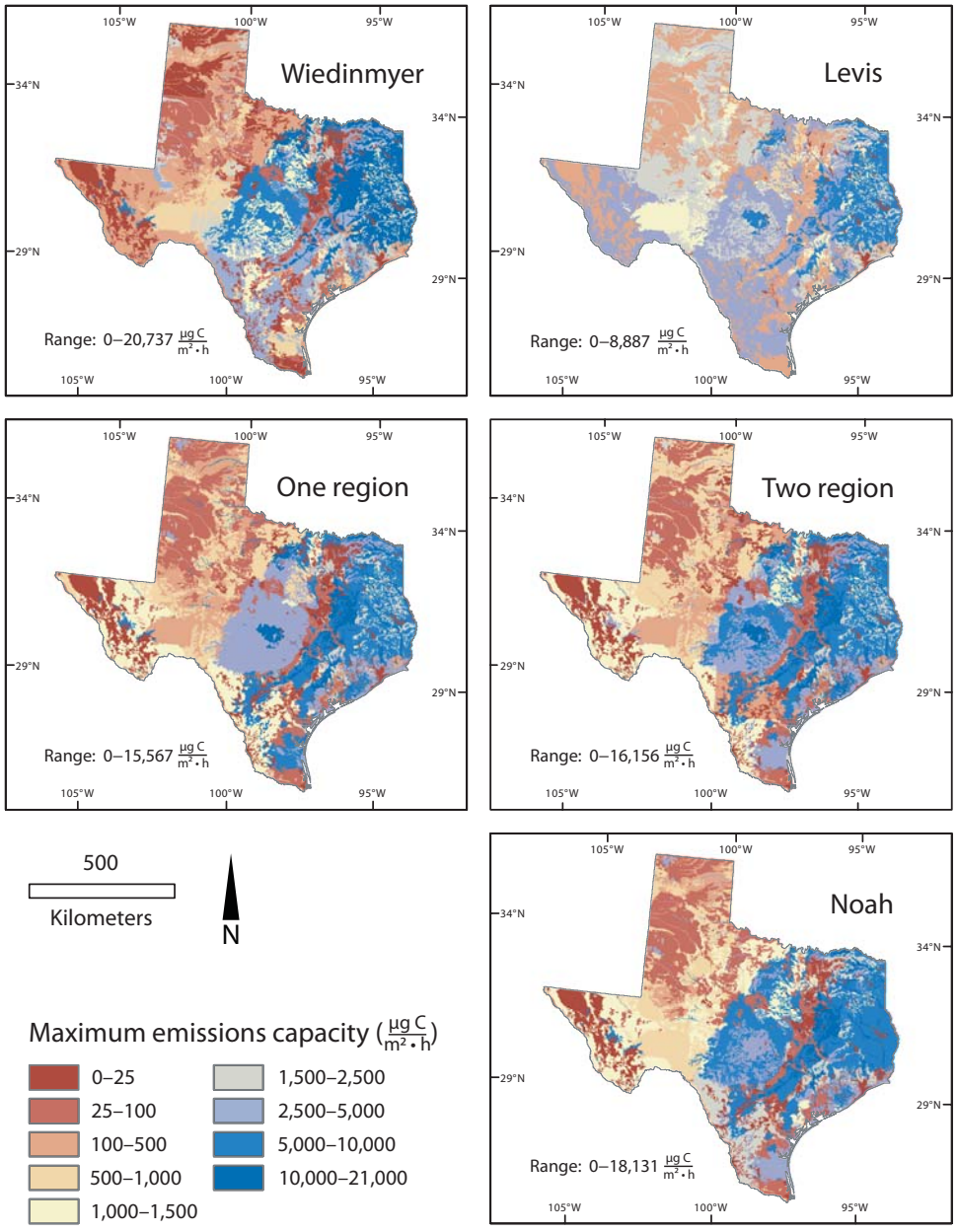

Figure 2.6. Total inherent BVOC emission capacity (sum of the inherent emission capacity for isoprene, monoterpene, and OVOC).

The panel labels (e.g., "Levis") describe the method used to calculate the total inherent BVOC emission capacity.

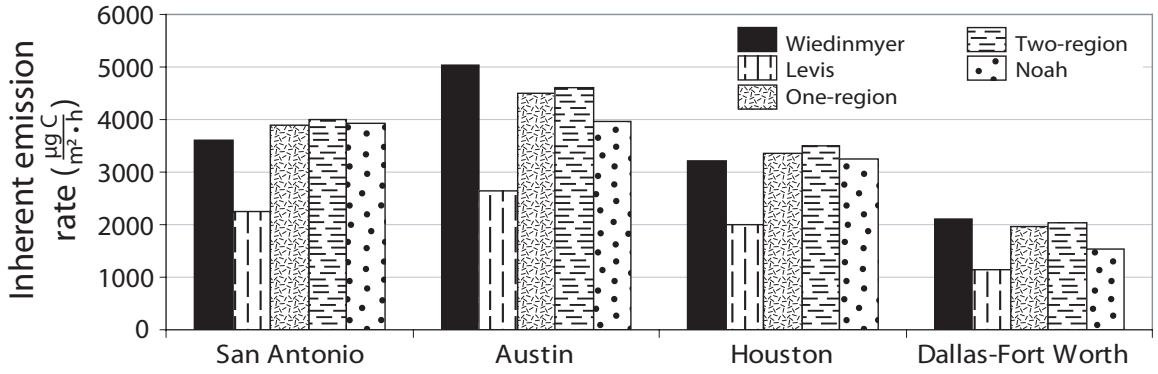

Figure 2.7. Comparison of the average inherent BVOC flux in four Texas cities.

Calculated using the "observed" species data ("Wiedinmyer"), the globally constant PFT-specific emission capacities used by Levis et al. ("Levis"), the Wiedinmyer-derived land-cover–type emission capacities for use in the Noah LSM ("Noah"), and the Wiedinmyer-derived PFT-specific emission capacities calculated using state-wide plant species data ("One-region") and species data from east and west Texas ("Two-region").

Chapter 3: Sensitivity of LSM-simulated biogenic emissions to the representation of land cover[2]

3.1. ABSTRACT

We evaluate the sensitivity of biogenic emissions simulated by a land-surface model (LSM) to different representations of land-cover vegetation. We drive the Community Land Model on a 0.1º grid over Texas, USA, from 1993–1998 using bilinearly interpolated North American Regional Reanalysis data. Two land-cover datasets provide the starting point for analysis: (1) a satellite-derived vegetation and soil color database and (2) a vegetation-distribution dataset derived from ground surveys. These datasets help us to qualitatively characterize the uncertainty in land-cover representations. We systematically vary the datasets to examine the sensitivity of modeled emissions to variation in representation of bare-soil fraction, vegetation-type distribution, and phenology.

Different datasets' representation of vegetation-type distribution leads to simulated mean statewide total biogenic emissions that vary by a factor of 3. Variation in specified bare-soil fraction causes simulated statewide average emissions that vary by a factor of 1.7. Scaling LAI values within reasonable bounds causes a near-linear change in simulated emissions. Differences in simulated values are largest for major metropolitan regions and for eastern and central Texas, where biogenic emissions are

[2]Substantial portions of this chapter were originally published in Gulden, L. E., Z.-L., Yang, and G.-Y. Niu, (2008a), Sensitivity of biogenic emissions simulated by a land-surface model to land-cover representations, *Atm. Env.*, doi:10.1016/j.atmosenv.2008.01.045. The References section contains full citations for all articles referenced here.

highest and where tropospheric ozone pollution is a significant concern. Changing bare-soil fraction alters simulated vegetation temperature and consequently indirectly affects modeled emissions (≤16% of inherent emissions capacity). Our estimates of model sensitivity to land-cover representation are consistent with those for other regions.

3.2. INTRODUCTION

Realistic simulation of biogenic emissions is a shared goal of climate scientists, the environmental engineering community, and air-quality policymakers. Biogenic volatile organic compounds (BVOCs) were first recognized as a key contributor to the formation of photochemical smog and were subsequently identified as actors in diverse climatic processes. In addition to their role in the production of tropospheric ozone, BVOCs condense to form secondary organic aerosols, which alter Earth's radiative balance (Kavouras et al., 1998; Claeys et al., 2004, Kroll et al., 2006) and serve as cloud condensation nuclei (Andreae and Crutzen, 1997). BVOCs are a non-negligible component of the global carbon cycle (e.g., Altshuller, 1991; Guenther, 2002).

Realistic representation of biogenic emissions within land-surface models (LSMs) is important for numerical weather forecasts, climate simulations, and the generation of accurate air quality forecasts. Because BVOCs transmit information about surface conditions between the land surface and the atmosphere, realistic simulation of the spatial distribution and magnitude of BVOC fluxes to the atmosphere likely improves the quality of simulated atmospheric conditions. Most research addressing how environmental change alters biogenic emissions and air quality has used meteorological model output to drive free-standing, task-specific biogenic emissions models and air-quality models (e.g.,

Tao et al., 2003). Synchronous coupling of consistent meteorological models, atmospheric chemistry models, and LSMs (e.g., Grell et al., 2006) has recently provided geoscientists with an improved capacity to simulate feedbacks between BVOC fluxes and other environmental components (Levis et al., 2003). LSM-simulated biogenic emissions such as those used within synchronously coupled systems inherit all sources of uncertainty associated with traditional biogenic emissions models (e.g., GloBEIS [Yarwood et al., 1999]) as well as additional sources of uncertainty unique to LSMs (e.g., coarse horizontal grid resolution and simplistic representation of vegetation [Gulden and Yang, 2006; Gulden et al., 2007a]).

Regardless of the type of model used, simulated BVOC flux estimates are notoriously uncertain. Guenther (1997) used six land-cover datasets to estimate the emission potential (units: μg C m^{-2} h^{-1} at 30ºC and 1000 μmol photons m^{-2} h^{-1}) of the contiguous United States and found a 3–5-fold difference in estimated inherent emission capacity. Other researchers, looking at sources of uncertainty ranging from meteorological inputs to vegetation-species distribution, have asserted that uncertainty in BVOC flux estimates ranges from a factor of 1.5 to more than 10, depending on the species of BVOC (e.g., Simpson et al., 1995; Simpson et al., 1999; Smiatek and Bogacki, 2005). Such uncertainty compromises air-quality simulations that depend on biogenic emissions estimates. Byun et al. (2005) used a satellite-derived dataset and the parent dataset of the ground-survey–derived dataset used here (Wiedinmyer et al., 2001) to show that selection of land-cover dataset can lead to a 10 ppb difference in simulated ozone concentrations for the metropolitan area surrounding Houston and Galveston, Texas, USA.

Although the true level of uncertainty in LSM land-cover datasets is difficult to quantify due to a dearth of observations (e.g., Simpson et al., 1999; Wiedinmyer et al., 2001), cursory examination of peer-reviewed literature introducing LSM-formatted land-cover datasets highlights the range of representations that have been deemed reasonable by the community (e.g., Bonan et al., 2002a; Wiedinmyer et al., 2001; Lawrence and Chase, 2006a).

In models that employ the Guenther et al. (1995) biogenic emissions algorithm or a close derivative (e.g., Levis et al., 2003), BVOC flux is a function of biomass, vegetation type, and environmental variation. A land-cover dataset's specified percentage of bare soil and the prescribed phenological variation determine the quantity of biomass on the modeled land-surface. The land-cover dataset also determines both the types and percent composition of vegetation types in a grid cell.

From all potential sources of uncertainty (Beck, 1987; Wagener and Gupta, 2005) in LSM-simulated biogenic emissions (e.g., lack of process understanding, oversimplification of physical processes, representation of heterogeneous vegetation with a limited number of plant functional types, model parameter uncertainty, uncertainty in meteorological forcing data, etc.), we focus here only on uncertainty that results from variation between datasets that specify land-cover characteristics. We perform a simple sensitivity analysis (Saltelli et al., 2000) to quantitatively and qualitatively attribute variation in LSM-simulated BVOC flux to different land-cover datasets.

We created several "cross-pollinated" land-cover datasets for the state of Texas, USA. Starting with a satellite-derived land-cover dataset (Lawrence and Chase, 2007) and a ground-survey–derived land-cover dataset (Wiedinmyer et al., 2001; Gulden and

Yang, 2006), we systematically varied, at each grid point in our model domain, the percent vegetated area, the vegetation distribution, and the magnitude of specified phenology. To assess the sensitivity of LSM-simulated biogenic emissions to the representation of surface vegetation, we used the hybrid datasets to initialize offline LSM runs. We examined the relative importance of the direct and indirect means by which divergent land-cover representations change simulated BVOCs.

We employ the National Center for Atmospheric Research's Community Land Model version 3 (CLM) (Oleson et al., 2004; Bonan et al., 2002b). CLM is a land-surface model (LSM) that incorporates a BVOC-flux module (Levis et al., 2003) founded on the work of Guenther et al. (1995). CLM is representative of LSMs commonly used in climate modeling. This research contributes to efforts to quantify the uncertainty associated with LSM-simulated biogenic emissions that is directly attributable to land-cover dataset (e.g., Guenther et al., 2006).

3.3. MODEL, METHODS, AND DATASETS

3.3.1. Representation of vegetation and biogenic emissions in CLM

CLM represents land cover as a mosaic of plant functional types (PFTs) (Bonan et al., 2002a). Each grid cell in the model domain is assigned a percent vegetated area; the vegetated area is further subdivided into four or fewer PFTs. The standard CLM, which uses static ecosystem dynamics, calculates daily variation in leaf area index (LAI) by linearly interpolating between prescribed monthly LAI values. Each PFT within model grid cell is assigned a unique LAI value for each month. Foliar density (the weight of dry leaf matter per unit area of ground covered by the PFT) is the product of LAI and specific

leaf area (m^2 leaf g^{-1} dry leaf matter). Although the static ecosystem dynamics of CLM allow LAI and stem area index (SAI) to vary seasonally, their seasonal cycle remains constant year to year.

CLM represents emission of isoprene, monoterpene, other volatile organic compounds, other reactive volatile organic compounds, and biogenic carbon monoxide (Levis et al., 2003; Guenther et al.,1995). For a given PFT, CLM simulates flux of BVOC type *i* as:

$$F_i = \varepsilon_i D \gamma_{PAR} \gamma_T \quad (3.1)$$

where F_i is the flux to the atmosphere of BVOC type *i* (units: μg C m^{-2} h^{-1}); D is the foliar density (units: g dry leaf matter (gdlm) m^{-2} of ground covered by the PFT), which is a scalar function of LAI; ε_i is a PFT-specific emission capacity for BVOC type *i* (units: μg C $gdlm^{-1}$ h^{-1}); γ_T is a dimensionless, nonlinear function of canopy temperature that modulates BVOC emissions; and γ_{PAR} is a dimensionless, nonlinear function of photosynthetically active radiation reaching the leaf surface that modulates isoprene emissions (for non-isoprene BVOCs, we assume $\gamma_{PAR} = 1$).

3.3.2. Starting-point land-cover datasets

Two source datasets provided the baseline for our analysis. The first is a 1-km PFT-distribution dataset developed from a ground-referenced, species-based dataset in which each grid cell contained between 1 and 115 of approximately 300 possible vegetation species (henceforth the "survey-derived dataset"). Wiedinmyer et al. (2001) describe the original species-based dataset; Gulden and Yang (2006) describe the conversion of the dataset to CLM format. The second raw dataset was derived from

Moderate Resolution Imaging Spectroradiometer (MODIS) and Advanced Very High Resolution Radiometer (AVHRR) satellite images (Lawrence and Chase, 2007) and contains 5-km resolution PFT distribution, plant phenological and structural parameters (LAI, SAI, height of canopy top and bottom), and soil-color information (henceforth the "satellite-derived dataset"). We interpolated both datasets to a uniform 0.1º grid. We used the two original datasets as parents from which we derived five unique, hybridized land-cover datasets. We used the hybridized datasets in lieu of the two originals such that we could more rigorously examine the sensitivity of the model to specific aspects of land-cover representation (i.e., PFT distribution, bare soil fraction, magnitude of specified phenology).

Figure 3.1 shows that the survey-derived dataset identifies 44.5% of the area of Texas as bare soil; the satellite-derived dataset labels 20.0% of Texas as bare soil. The difference in bare soil fraction is especially pronounced in central and eastern Texas. Trees cover 20.3% of Texas in the survey-derived dataset; in the satellite-derived dataset trees cover 10.7%. Introduced in the peer-reviewed literature, both datasets have been deemed reasonable representations of reality by the scientific community and are used for both scientific and engineering purposes (P. Lawrence, personal communication; e.g., Junquera et al., 2005). Variations in simulated biogenic emissions that result from the use of these two datasets help us to qualitatively characterize the uncertainty in LSM-simulated emissions that stem from uncertain land-cover representations.

3.3.3. Modified datasets

The original survey-derived dataset specifies only PFT fraction; it contains no phenological information. Before creating the hybrid datasets, we transformed the satellite-derived phenological parameters for use with the survey-derived PFT fractions. For a given grid cell, the satellite-derived dataset contains LAI and stem area index (SAI) information only for the PFTs identified as being present in that location. Because the specified vegetation distributions differ between the satellite- and survey-derived datasets, we could not directly combine the satellite-derived LAI and SAI with the survey-derived PFT distribution data. To ensure that LAI and SAI were defined at all model grid points for all PFTs (as required by CLM), we used the area-weighted, longitudinal average of the satellite-derived PFT parameters. Vegetation biomass in the state of Texas exhibits a very large west–east gradient: an area-weighted average along lines of longitude was deemed more appropriate than a statewide average, a latitudinal average, or another averaging method.

The satellite-derived dataset does not identify broadleaf evergreen trees (BETs) or broadleaf evergreen shrubs (BESs) in Texas; it contains neither LAI nor SAI information for BET and BES. We defined the LAI and SAI of BET for all months as the June–July–August (JJA) average of the longitudinally averaged LAI and SAI for broadleaf deciduous trees. We defined LAI and SAI for BES for all months as the JJA average of broadleaf deciduous shrubs. Needleleaf evergreen shrubs (NESs) and temperate needleleaf deciduous trees (NDTs) do not exist in the standard CLM, but they do exist in the survey-derived dataset (Gulden and Yang [2006]). We defined the LAI and SAI of NES as identical to those of BES. LAI and SAI for NDT were defined using the LAI and

SAI of needleleaf evergreen tree. Because there are so few NDTs in Texas, we neglected any error caused by spuriously high NDT biomass in wintertime. Our method for defining LAI and SAI for each PFT, for each site, and for each month for the survey-based dataset did not introduce significant error: statewide BVOC flux estimates derived using the dataset with the satellite-derived PFTs, satellite-derived bare soil, and the longitudinally averaged, satellite-derived phenological parameters were consistently within 1% of estimates derived using the original satellite-derived dataset (results not shown).

We created three pairs of hybrid datasets, designing them to facilitate a clean comparison of the effect on biogenic emissions of two different realizations of a single descriptor of land-surface vegetation.

3.3.3.1. Dataset pair 1: PFT-SURVEY and PFT-SATELLITE

The fraction of the vegetated area covered by each PFT controls the spatial distribution and magnitude of a dataset's inherent BVOC flux: trees emit BVOCs at a rate often an order of magnitude greater than the emission rate of grasses and crops. The first pair of hybrid datasets isolated the effect of uncertainty in PFT distribution on simulated BVOC flux. The pair differed only in its PFT distribution. Derivative dataset PFT-SURVEY represents vegetation composition with the survey-derived PFT distribution. PFT-SATELLITE uses the satellite-derived PFT distribution. Both use the satellite-derived, longitudinally averaged phenological parameters and the percent bare soil specified in the satellite-derived dataset.

3.3.3.2. Dataset pair 2: BARE-SURVEY and BARE-SATELLITE

The partitioning of a grid cell between vegetated area and bare soil helps determine a grid cell's biomass density and affects the modeled energy and water balances. When all else is equal, a greater amount of biomass increases BVOC flux. The second set of derivative datasets allowed us to examine the sensitivity of modeled biogenic emissions to the range of realistic assessments of bare soil fraction. The pair differed only in its bare soil fraction. Dataset BARE-SURVEY uses the percent bare soil specified in the survey-derived dataset; BARE-SATELLITE uses the percent bare soil specified by the satellite-derived dataset. Both use the satellite-derived, longitudinally averaged phenological parameters; both use the survey-derived PFT distribution. (It is important to note that, although their bare soil fraction differs, the percent of the vegetated area covered by a given PFT remains constant between the two datasets.) BARE-SATELLITE is more densely vegetated than BARE-SURVEY, especially in central and eastern Texas. Note that BARE-SATELLITE and PFT-SURVEY are identical.

3.3.3.3. Dataset pair 3: LAI×0.5 and LAI×1.5

CLM calculates BVOC flux as a linear function of biomass density, which is a scalar multiple the PFT's LAI. To simulate uncertainty in the magnitude of phenological parameters, we modified the LAI and SAI values in the original satellite-derived dataset to create artificial datasets with underestimated and overestimated LAI and SAI values. We uniformly scaled the satellite-derived LAI and SAI values by 0.5 to create derived dataset LAI×0.5 and multiplied the satellite-derived LAI and SAI values by 1.5 to create

LAI×1.5. Both datasets used the satellite-derived PFT distribution and the satellite-derived vegetation fraction information. LAI×0.5 approximates the lower bound of realistic LAI values, and LAI×1.5 is an estimate of a reasonable upper bound for realistic LAI values. These bounds are consistent with the range of LAI values presented in the literature (e.g., Tian et al., 2004). Because the satellite-derived PFT distribution was used for all three datasets, the phenological parameters in both LAI×0.5 and LAI×1.5 were taken directly from the original satellite-derived phenological data (i.e., they were not longitudinally averaged).

Table 3.1 summarizes the characteristics of the three pairs of datasets. Hereafter, model runs will be referred to according to their land-surface dataset name (e.g., "PFT-SATELLITE" will be used to mean "the model run that employed PFT-SATELLITE as its input land-cover dataset").

3.3.4. Parameters for model runs

Meteorological input forcing was obtained from the North American Regional Reanalysis (NARR) data (Mesinger et al., 2006). We used bilinear interpolation to convert the NARR data from their original 32-km grid to a 0.1º grid coincident with the land-surface datasets. All runs represented the period from January 1, 1993, to January 1, 1999. We analyzed model output from January 1, 1995–January 1, 1999, a period spanning both a weak La Niña event and a strong El Niño event.

For all experiments, we used region-specific, PFT-specific emissions capacities. The emissions capacities were developed using a ground-referenced, species-based dataset (Wiedinmyer et al., 2001) and its CLM-compatible counterpart (Gulden and

Yang, 2006). The CLM-compatible dataset is the survey-based dataset used as one of the two "raw" datasets.

For all runs, we employed static ecosystem dynamics and used CLM-standard values for top and bottom heights of the canopy, both of which are spatially and temporally constant and vary only between PFTs. The satellite-derived dataset provided soil color for all simulations; soil texture was defined using the 5-minute CLM-standard soil texture data.

3.4. RESULTS

As expected, LSM-simulated biogenic emissions are sensitive to land cover: significant variation in model output can be directly attributed to uncertainty in land-cover dataset. We use the mean scale factor to measure change in model response that is directly attributable to a change in a specific vegetation parameter (e.g., PFT distribution). The scale factor for a given grid cell is time average ratio of BVOC flux estimates for a pair of runs. The mean scale factor is the domain average scale factor. Table 3.2 summarizes the results described in the following sections (3.4.1–3.4.3) and provides additional information regarding regional differences in scale factors.

3.4.1. Sensitivity to representation of the distribution of PFTs

Estimates of the statewide mean monthly BVOC flux derived from PFT-SURVEY were on average 3 times as great as those derived by PFT-SATELLITE (Figure 3.2) (i.e., the mean scale factor is 3.0). This difference is greatest during summer, when the difference between tree biomass and herbaceous biomass is greatest.

Figure 3.3 shows the spatial distribution of the differences in monthly mean June–July–August (JJA) BVOC emissions, averaged over the period of analysis. The mean scale factor varies considerably between regions: in west Texas, where biogenic emissions are relatively low, the mean scale factor is 2; in central Texas (including San Antonio and Austin) it is ~6.

Because the survey-derived dataset was created using a suite of smaller, ground-referenced datasets, the survey-derived PFT distribution (used in PFT-SURVEY) may better represent reality than does the PFT distribution used in PFT-SATELLITE. However, the survey-derived dataset likely overestimates the percentage of trees in central Texas (Gulden and Yang, 2006). "Reality" probably lies somewhere between the two representations.

3.4.2. Sensitivity to representation of the fraction of bare soil

The statewide total BVOC flux simulated by BARE-SATELLITE is on average 1.7 times as large as that simulated by BARE-SURVEY. Figure 3.4 provides the time series of monthly statewide total BVOC flux over the period of analysis.

Figure 3.5 shows the differences in the spatial distribution of the mean JJA BVOC flux over the study period. The mean scale factor associated with bare-soil fraction does not vary by region as much as does the mean scale factor associated with PFT distribution; however, the Dallas–Fort Worth metropolitan area has a considerably higher uncertainty associated with bare-soil fraction than do other regions (mean scale factor = 3.3).

The satellite data gathers bare-soil fraction from a near-nadir perspective and is therefore arguably more in agreement with the "viewpoint" of the LSM. It is reasonable to assume that the bare-soil fraction derived from the satellite images is closer to reality, but this assertion is difficult to validate.

3.4.3. Sensitivity to magnitude of specified phenology

As biomass increases, light-activated biogenic emissions are expected to decrease in regions of the canopy that become increasingly shaded. However, when LAI increases, a larger area of leaves receives solar radiation. When shading effects are combined with nonlinear canopy-temperature controls on emissions, it is not clear whether increasing biomass causes a corresponding linear increase in biogenic emissions.

CLM represents canopy temperature (as functions of sensible and latent heating and cooling processes) and canopy shading processes (as a function of LAI and SAI), both of which may vary nonlinearly as LAI increases. Biogenic emissions in CLM depend on environmental modulation factors γ_{PAR} and γ_T, which are nonlinear functions of radiation reaching the canopy and canopy temperature (Figure 3.6). Because of these process representations, CLM has the potential to help us identify whether changing canopy shading and changing canopy temperature are significant controls on biogenic emissions as LAI varies.

The mean scale factor for the dataset pair made up of the LAI×1.5 and the "raw" original satellite-derived dataset is ~1.5. The baseline dataset produces an estimate that is on average ~2 times that of LAI×0.5. Figure 3.7 compares estimates obtained from LAI×0.5 and LAI×1.5 to the baseline run. Regional differences in mean scale factor are

negligible (Table 3.2). The linear functional relationship between model biomass and modeled BVOC flux (Equation 3.1) causes a coincident, nearly linear scaling of model-simulated BVOC flux when LAI and SAI are scaled linearly. In the model (if not in nature), biogenic emissions are more sensitive to changes in biomass than to consequent changes in canopy shaded fraction or canopy temperature.

In the case of LAI, indirect effects on biogenic emissions modulated by environmental modulation factors γ_T and γ_{PAR} are relatively insignificant. Monthly average values of γ_T remained stable; γ_{PAR} did change slightly between LAI runs, but that change was not significant (results not shown). Whether CLM adequately represents canopy temperature and canopy shading processes as functions of changing LAI remains an open research question: to our knowledge, relatively little evaluation of these parameterizations has been done.

3.4.4. Indirect effect of land-cover representation on biogenic emissions

Bare soil fraction, the distribution of PFTs, and phenological parameters indirectly control on the actual CLM3-simulated BVOC flux; however, the magnitude of each effect varies. Figure 3.8a shows the time series of γ_T for isoprene for BARE-SURVEY and BARE-SATELLITE; Figure 3.8b shows the corresponding time series for γ_T for nonisoprene BVOCs. Bare soil fraction has the greatest impact on modeled leaf-surface temperature (and hence on γ_T). The higher mean leaf-surface temperature simulated by BARE-SATELLITE caused an increase in total modeled isoprene flux of up to 5% of inherent BVOC flux when compared to BARE-SURVEY. The difference in simulated vegetation temperature was responsible for an increased total non-isoprene flux

of up to 16% of the inherent BVOC flux. This is particularly important in the needleleaf evergreen forests of eastern Texas, where nonmethane biogenic hydrocarbon emissions are, at least in the model, dominated by nonisoprene BVOCs. Whether such indirect variation is representative of natural processes remains an open question.

The difference between the γ_T modulation factors for both isoprene and nonisoprene-BVOCs that were calculated by LAI×0.5 and LAI×1.5 was in all cases less than 1% (results not shown). PFT-SURVEY tended to simulate slightly higher vegetation temperatures that PFT-SATELLITE. Model γ_T values were correspondingly higher: PFT-SATELLITE γ_T for nonisoprene BVOCs ranged from 1% greater to 4% less than the corresponding γ_T calculated by PFT-SURVEY. Differences between γ_T for isoprene between PFT-SATELLITE and PFT-SURVEY were between 1% and 2%.

3.5. DISCUSSION

The representation of vegetation characteristics is by no means the only source of uncertainty in LSM-simulated biogenic emissions. Even if a land-cover dataset perfectly describes the "true" land-cover distribution and biomass density of a landscape, in nature, spatial variation of the species composition of can vastly alter the landscape's inherent biogenic emission flux (e.g., Guenther et al., 1994; Guenther, 1997). LSMs, including CLM, may fail to accurately simulate canopy temperature or the PAR reaching the leaf surface. Even if LSMs were to represent the "true" grid-cell average state variables, the nonlinearity of the response of BVOCs to environmental variation makes model error sensitive to subgrid-scale variation. Error in leaf-level species-based emission capacity measurements, interspecies and intraspecies variation (e.g., Funk et al., 2005) in biogenic

emissions, the unproven universal applicability of the Guenther et al. (1995) algorithm (e.g., Schuh et al, 1997), and the omission from the CLM version of the emission algorithm of factors important to emissions (such as leaf age [e.g., Zhang et al., 2000, Guenther et al., 2006]) all contribute additional uncertainty to LSM-generated BVOC fluxes. Consequently, any projection of future biogenic emissions derived from LSMs, either offline or coupled to climate models, should be viewed only as a very rough estimate.

The statewide total BVOC flux estimates that results from variation in the three vegetation descriptors examined varies by up to a factor of 3; however, the range of estimates varies significantly by region. Sensitivity of simulated emissions to use of different land-cover datasets is highest in regions where BVOC emissions are of greatest concern: mean scale factors are largest in central and eastern Texas (east of –99.5ºE), where wooded and forested landscapes ensure that BVOC flux is relatively high and where tropospheric ozone pollution is a primary concern for air-quality managers. It is also worth noting that a mean scale factor of 3.2 in eastern Texas corresponds to a much larger actual difference in the mass of BVOCs emitted than would a mean scale factor of 3.2 in western Texas, where biomass density is low and tree cover is sparse. On first-order examination, range of simulated emissions values is particularly large in the major metropolitan regions of the state (see Table 3.2).

The largest source of land-cover–related variation in LSM-simulated biogenic emissions is PFT distribution. BVOC emissions estimates within LSMs and their coupled models of the atmosphere will benefit if future land-cover data acquisition projects focus their efforts on precisely quantifying the bare-soil fraction and on the partitioning of the

vegetated area between trees and grass. Of secondary—but certainly non-negligible—importance is to improve our confidence in observed estimates of the quantity of biomass present on the ground, which is determined by the percent vegetated area and by the magnitude and seasonality of leaf phenology, the latter of which we did not examine here (see Gulden et al., 2007a).

We consider only the monthly mean variation in BVOC flux, emission modulation factors, and model state variables. Divergent representations of the land surface may significantly alter estimates of the diurnal cycles of emission modulation factors, canopy state variables, and, consequently, BVOC emissions. Examination of uncertainty at a finer temporal resolution is warranted.

Use of static ecosystem dynamics considerably underestimates the true variation in biogenic emissions (Gulden et al., 2007a). When employing a dynamic vegetation module that updates changes in maximum LAI once a year and allows daily variation in the fraction of the maximum LAI in response to environmental conditions, Levis et al. (2003) found that interannual variation in total BVOC flux exceeded 29% during a 10-year fully coupled climate simulation. Gulden et al. (2007a) showed that, the absolute average departure from the monthly mean (max) BVOC flux was 22.4% (137%) when phenology was allowed to respond to short-term environmental variation.

Our estimates of the sensitivity of simulated biogenic emissions, which focus on only one source of potential error in LSM-simulated BVOC flux and which we derived from observations-based land-cover datasets, are of the same order of magnitude as previous estimates (e.g., Simpson et al., 1999; Hanna et al., 2005).

3.6. SUMMARY AND CONCLUSIONS

We conclude: (1) Differences in the representation of land-surface vegetation in the state of Texas, USA, can result in estimated monthly mean total biogenic emissions that differ by up to a factor of 3. (2) Distribution of PFTs contributes most to this variation. The ground-survey–derived PFT distribution resulted in statewide monthly mean BVOC fluxes that were an average of 3 times as large as the estimates produced by a run using a satellite-derived PFT distribution. (3) Divergent representations of bare-soil fraction significantly contribute to variation in simulated BVOC flux: a run using satellite-derived bare soil fraction produced BVOC estimates 1.7 times as great as a run using the bare-soil fraction from the less-densely-vegetated ground-survey–derived dataset. (4) Scaling LAI within reasonable bounds (50–150% of the satellite–derived estimates) caused a nearly linear decrease and increase, respectively, of simulated biogenic emissions. (5) Sensitivity to land-cover dataset is highest in central and eastern Texas (east of –99.5ºE), where there is up to a 6.3-fold difference between the modeled BVOC flux when different datasets are used. (6) Variation between emissions estimates is especially pronounced in major metropolitan areas, where ozone pollution significantly degrades urban air quality. (7) Different specifications of bare soil fraction can have a significant indirect effect on modeled actual BVOC flux (up to 16% of inherent BVOC flux) through modification of state variables that control vegetation temperature; however, we do not know whether the modeled indirect effects are model artifacts or representations of reality.

Urban planners and air quality managers who use LSM-based model predictions of BVOC emissions should be aware of the sensitivity of modeled BVOC flux estimates

to uncertainty in the land-cover dataset used. When LSMs are linked to climate models, this sensitivity may propagate uncertainty to all BVOC-related radiative, carbon-cycle, and atmospheric-chemistry processes. Although our results specifically address the simulation of biogenic emissions within LSMs, they apply broadly to any application of an LSM in which the variable of interest depends in part on land-cover dataset used.

3.7. ACKNOWLEDGEMENTS

We thank Peter Lawrence, who provided us with the satellite-derived land-cover dataset and who lent us many useful insights. We are also grateful for the insightful input of Enrique Rosero and for additional constructive comments made by two anonymous reviewers, Alan Robock, and Mark Estes. This work was funded by the U.S. Environmental Protection Agency's Science To Achieve Results (STAR) Program Grant #RD83145201 and by the National Science Foundation. We thank the Texas Advanced Computing Center for computing resources.

Table 3.1. Source data used to create derivative dataset pairs.

	Pair 1		Pair 2		Pair 3	
Dataset name	PFT-SURVEY	PFT-SATELLITE	BARE-SURVEY	BARE-SATELLITE	LAI×1.5	LAI×0.5
PFT distribution	Survey-derived	Satellite-derived	Survey-derived	Survey-derived	Satellite-derived	Satellite-derived
Bare soil fraction	Satellite-derived	Satellite-derived	Survey-derived	Satellite-derived	Satellite-derived	Satellite-derived
Phenology	Satellite-derived; longitudinal averaged	Satellite-derived; longitudinal averaged	Satellite-derived; longitudinal averaged	Satellite-derived; longitudinal averaged	Original satellite-derived, uniformly scaled by 1.5	Original satellite-derived, uniformly scaled by 0.5

Table 3.2. Regional variation in the mean emission scale factor.*

		Pairs of land-cover datasets**			
		PFT-SURVEY / PFT-SATELLITE	BARE-SATELLITE / BARE-SURVEY	LAI×1.5 / Baseline satellite-derived	Baseline satellite-derived / LAI×0.5
Entire state of Texas	Mean	3	1.7	1.5	2
	Range	2.6–3.2	1.6–1.7	1.5–1.5	1.9–2.0
East Texas (east of –99.5ºE)	Mean	3.2	1.7	1.5	2
	Range	2.7–3.5	1.6–1.7	1.5	1.9–2.0
West Texas (west of –99.5ºE)	Mean	2	1.7	1.5	2
	Range	1.7–2.3	1.6–1.7	1.4–1.5	1.9–2.1
Houston§	Mean	3.3	1.9	1.4	2
	Range	2.4–3.8	1.8–1.9	1.4–1.5	1.9–2.0
Dallas/Fort Worth†	Mean	4.6	3.3	1.5	2
	Range	3.5–5.3	3.1–3.5	1.4–1.5	1.9–2.0
Central Texas (San Antonio, Austin)‡	Mean	5.9	1.9	1.5	2
	Range	5.1–6.3	1.8–1.9	1.5	1.9–2.0

* The "scale factor" is the regional average ratio of BVOC flux estimates for a pair of runs. Monthly mean total BVOC flux estimates were calculated for each 0.1º×0.1º model grid cell for each month of the simulation period (1995–1998). For instance, in the PFT-SURVEY/PFT-SATELLITE column, the model run that used the PFT-SURVEY dataset produced monthly mean flux estimates that were on average 3.0 times as great as those produced by the run using the PFT-SATELLITE data.

****** Column headings are the land-cover datasets used in each of the two paired runs.

§ Houston metropolitan area defined as –94.6ºE to –96.0ºE and 29.15ºN to 30.35ºN

† Dallas/Fort Worth defined as 32.40ºN to 33.40ºN and –96.35ºE to –97.7ºE.

‡ "Central Texas" (spanning both the approximate Austin and San Antonio metropolitan areas) defined as the land between –97.45ºE to –99.0ºE and 29.15ºN to 30.75ºN.

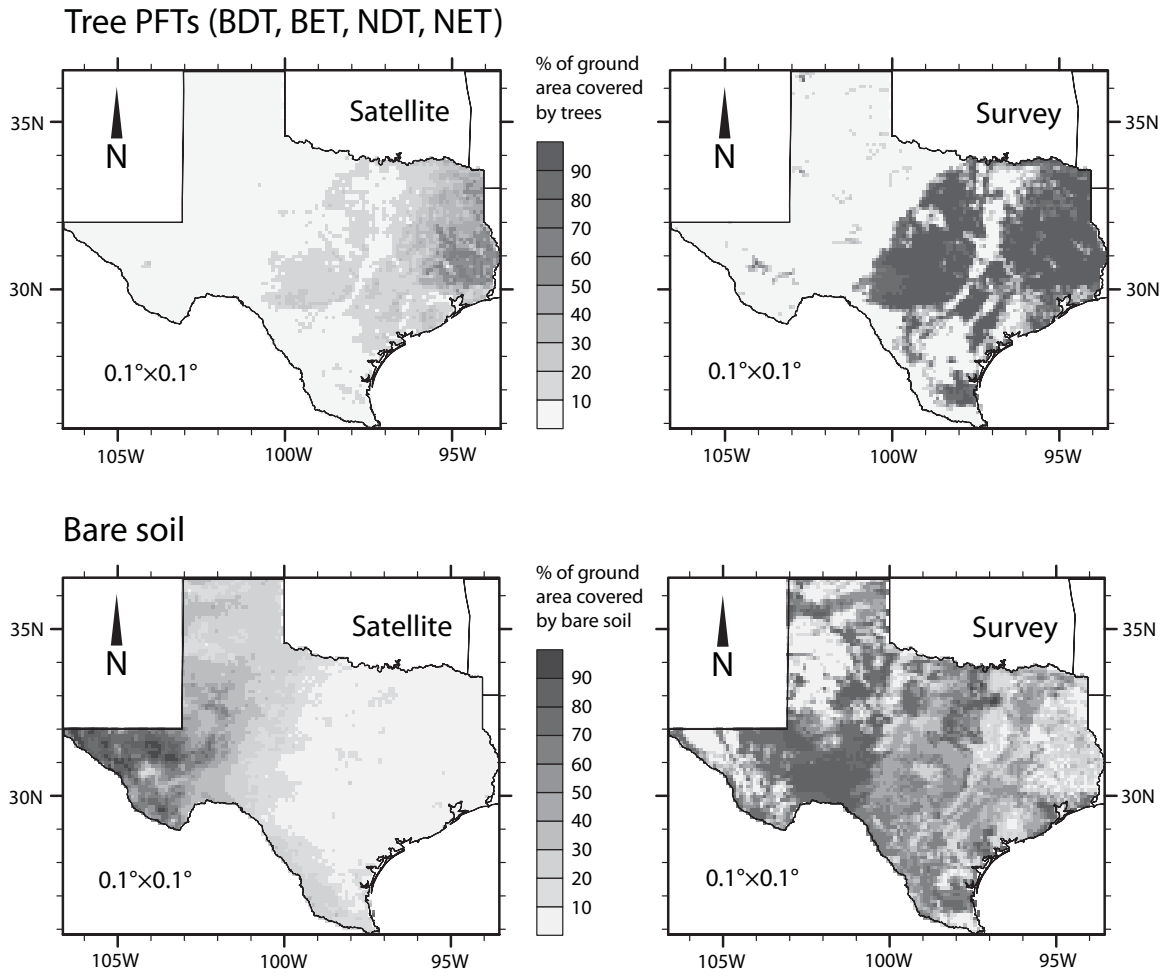

Figure 3.1. Summary of the biogenic-emissions–relevant differences between the satellite-derived land-cover dataset and the survey-derived land-cover dataset.

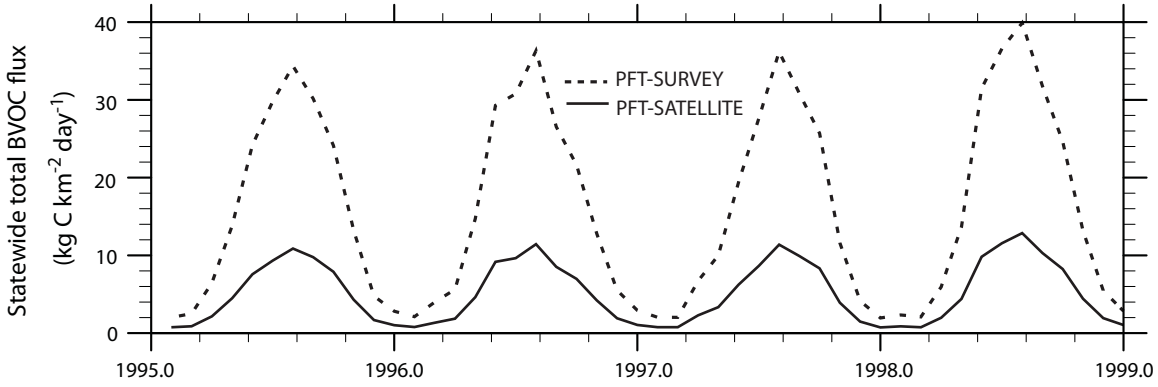

Figure 3.2. Time series of BVOC flux generated by the run using PFT-SATELLITE and the run using PFT-SURVEY.

Flux is the monthly average, statewide total flux.

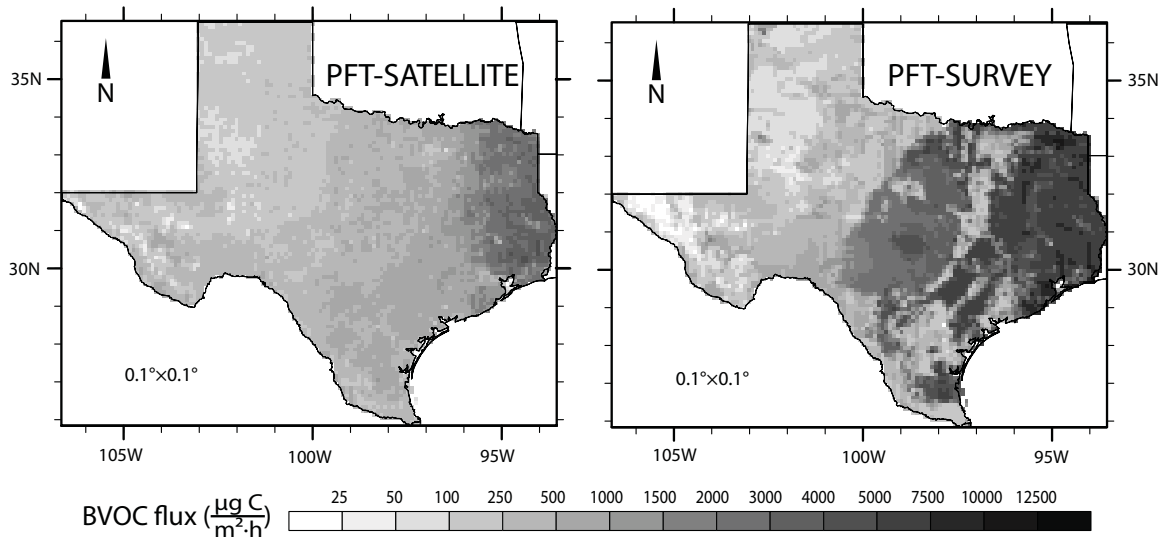

Figure 3.3. Comparison of BVOC flux estimates from PFT-SATELLITE and PFT-SURVEY.

June–July–August mean BVOC flux, averaged over the analysis period (1995–1998).

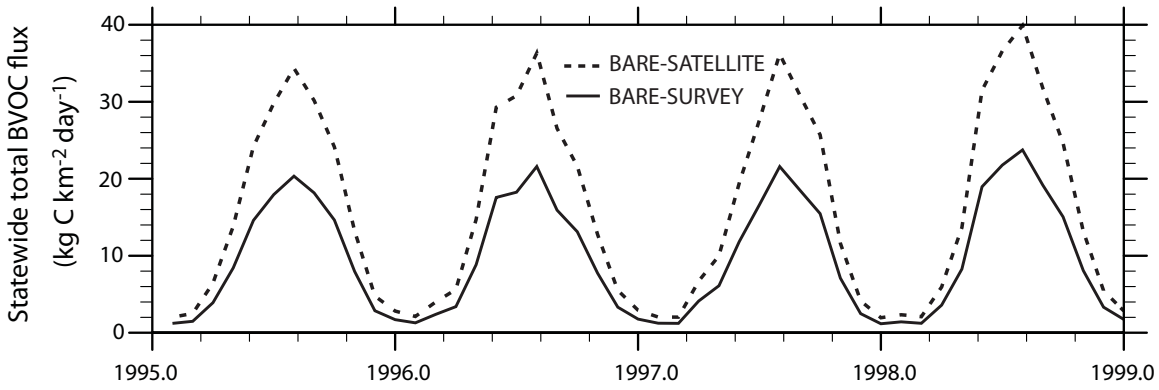

Figure 3.4. Time series of BVOC flux simulated by the run using BARE-SATELLITE and by the run using BARE-SURVEY.

Flux is the monthly average, statewide total flux.

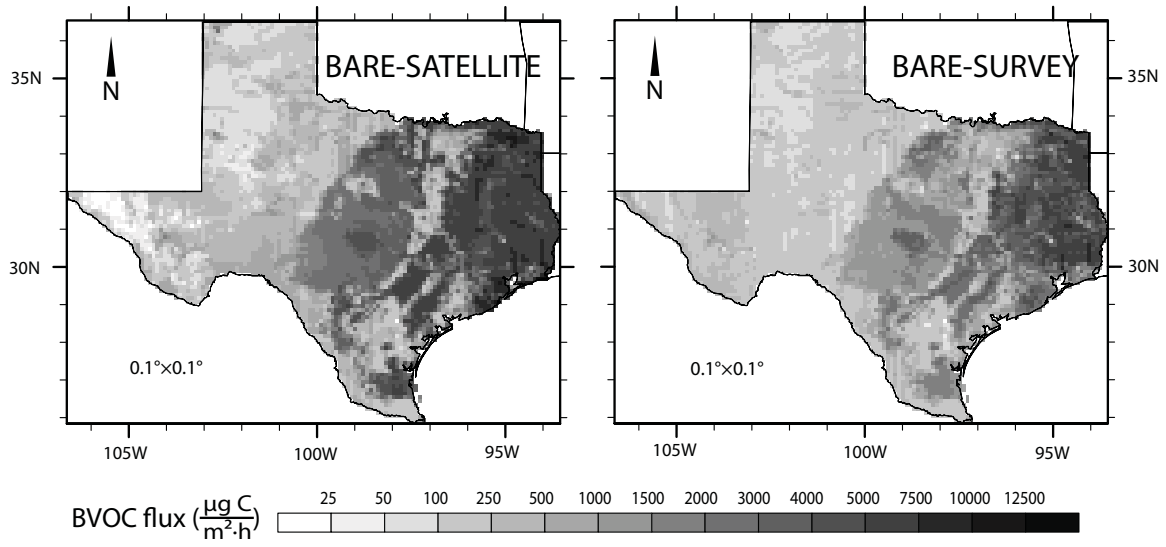

Figure 3.5. Comparison of flux estimates from BARE-SATELLITE and BARE-SURVEY.

June–July–August mean BVOC flux, averaged over the analysis period (1995–1998).

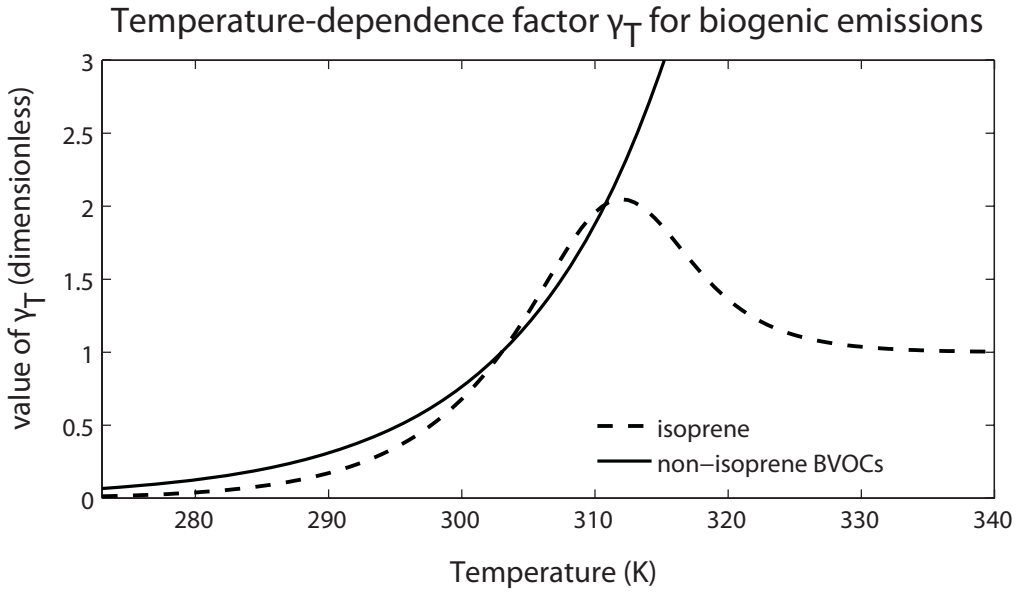

Figure 3.6. Functions governing the environmental modulation factor γ_T for isoprene and non-isoprene BVOCs.

γ_T allows CLM-simulated biogenic emissions to respond to variations in leaf temperature (Equation 3.1).

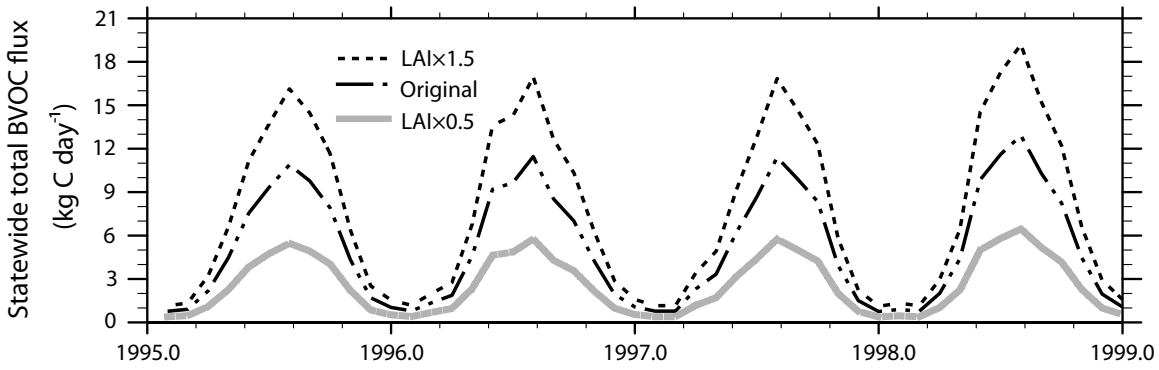

Figure 3.7. Time series of BVOC flux generated by the run using LAI×0.5, the original satellite-derived land-cover dataset, and LAI×1.5.

Flux is the monthly average, statewide total flux.

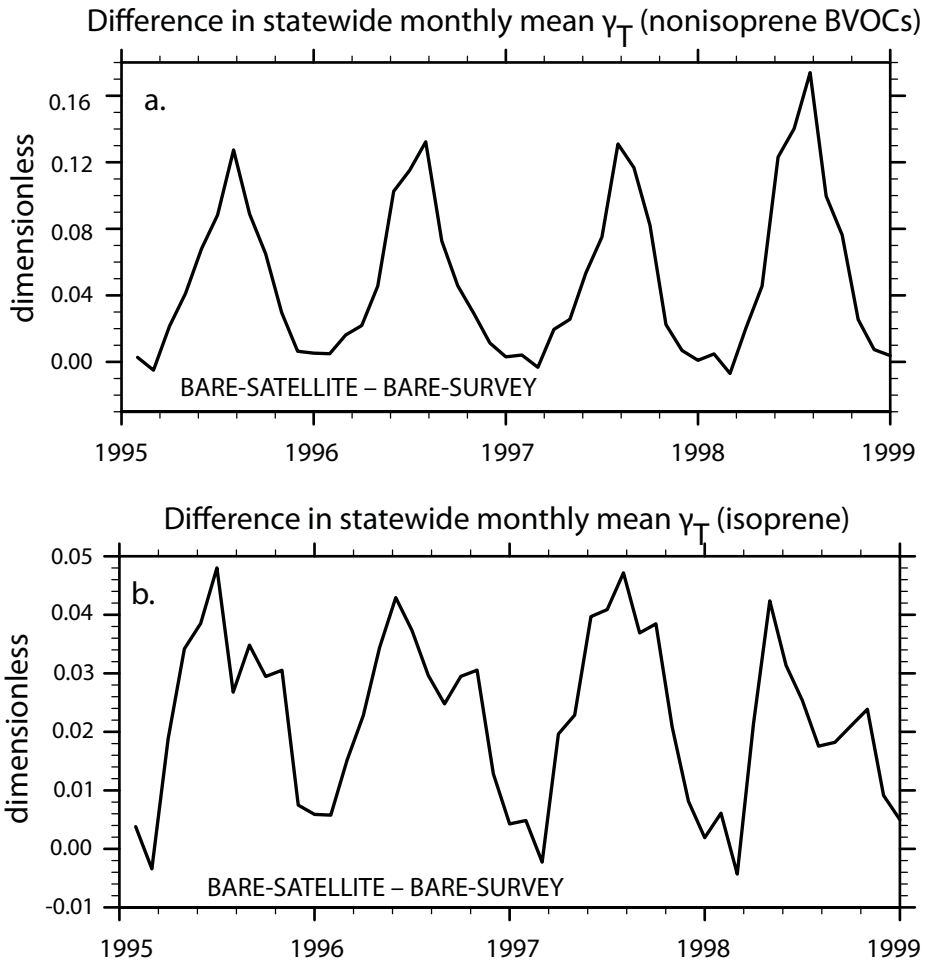

Figure 3.8. Comparison of the leaf-temperature emission-modulation factors calculated by BARE-SURVEY and BARE-SATELLITE for both isoprene and non-isoprene BVOCs.

Both panels show BARE-SATELLITE less BARE-SURVEY. On average, leaf temperature simulated by BARE-SATELLITE was higher than that simulated by BARE-SURVEY.

Chapter 4: Year-to-year variation in biogenic emissions on a regional scale[3]

4.1. ABSTRACT

Especially on regional scales, interannual variation in biogenic emissions is not well quantified. We use a land-surface model augmented with a short-term dynamic phenology scheme to estimate the interannual variation in the emission of biogenic volatile organic compounds (BVOCs) between 1982 and 2004.

We use North American Regional Reanalysis data to drive two versions of the National Center for Atmospheric Research Community Land Model (CLM) on a 0.1º grid over eastern Texas. The first version is the standard CLM with prescribed leaf area index (LAI) (i.e., LAI varies seasonally but not interannually); the second version is the standard CLM augmented with a dynamic phenology scheme (CLM-DP) that allows LAI to respond to environmental variation. We calibrate CLM-DP using satellite-derived LAI as our visual constraint.

When phenology is prescribed, the domain-mean (domain-maximum) average absolute departure from the monthly mean BVOC flux is 11.7% (70.6%); when phenology is allowed to vary with environmental conditions, it is 22.4% (137.7%). The domain-mean (domain-maximum) average absolute departure from the monthly mean flux is lower during summer: using CLM-DP, it is 15.7% (35.3%); using the standard

[3]Substantial portions of this chapter were originally published in Gulden, L. E., Z.-L. Yang, and G.-Y. Niu (2007a), Interannual variation in biogenic emissions on a regional scale, *J. Geophys Res.*, 112, D14103, doi:10.1029/2006JD008231. The References section contains full citations for all articles referenced here.

CLM, it is 7.0% (23.0%). The domain-average, mean-normalized standard deviation of the June–July–August mean BVOC flux is 0.0619 when LAI is prescribed and 0.183 when LAI varies with environmental conditions.

Our results imply that interannual variation of leaf-biomass density, which is primarily driven by interannual variability of precipitation, is a significant contributor to year-to-year differences in BVOC flux on a regional scale, of at least equal importance to interannual variation of temperature and shortwave radiation. Phenology-driven biogenic emission variability is most pronounced in regions with relatively low emissions: as a grid cell's mean BVOC flux decreases, the mean-normalized standard deviation of BVOC flux tends to increase. BVOC flux is most variable between years in sub-humid, sparsely wooded regions where interannual variability of precipitation is relatively large.

4.2. INTRODUCTION

Biogenic volatile organic compounds (BVOCs) are involved in a suite of environmental processes. BVOCs condense to form secondary organic aerosols (Kavouras et al., 1998; Claeys et al., 2004), which often become cloud condensation nuclei (Novakov and Penner, 1993) and which alter the radiative balance at Earth's surface (Andreae and Crutzen, 1997). BVOCs react in the presence of nitrogen oxides to increase the concentration of tropospheric ozone, which is a respiratory irritant and a primary constituent of photochemical smog (Chameides et al., 1988). High concentrations of ozone adversely affect plant photosynthesis and growth (Ashmore, 2005). The ultimate reaction product of most BVOCs is carbon dioxide (Guenther, 2002), and the presence of isoprene may increase the lifetime of methane (Poisson et al., 2000),

a potent greenhouse gas. Because a diverse set of environmental processes responds to forcing by biogenic emissions, year-to-year variation in the flux of BVOCs has the potential to influence interannual climate variability.

Many factors affect the rate of biogenic emissions. The vegetation species composition of a landscape—which is controlled in large part by local climate—exerts primary control on the flux of BVOCs from the land surface. The rate of biogenic emissions also has been shown to vary as a function of leaf biomass density, the amount of photosynthetically active radiation (PAR) reaching the leaf surface, the canopy temperature, leaf age (Guenther et al., 1991; Monson et al., 1994; Fuentes et al. 1995), soil nutrient availability, ambient carbon dioxide concentration (Possell et al., 2004), drought stress, and the leaf-to-air vapor-pressure deficit (Pegoraro et al., 2005). Multiple researchers have examined factors controlling seasonal variation in biogenic emissions (e.g., Monson et al., 1994; Fuentes et al., 1999). Here we focus on interannual variation.

Interannual variation in BVOC flux is likely considerable. Levis et al. (2003) used a biogenic emissions module (Guenther et al., 1995) and a dynamic global vegetation module (DGVM) within a climate model to show that the interannual variation in the total global flux of BVOCs in a given month exceeded 18% during a 10-year fully coupled climate simulation. Naik et al. (2004), drove the Integrated Biosphere Simulator (IBIS), a DGVM, from 1971–1990 and found that modeled seasonal variability of total global biogenic emissions ranged from 17–25%. Tao and Jain (2005) used a terrestrial biosphere model to represent 1981–2000 and found that, when they represented both carbon dioxide fertilization of the biosphere and monthly scale climate variability, global total isoprene emissions for a given month varied up to 31% between years. These

estimates of interannual variation in monthly BVOC flux are lower than the satellite-derived estimate of 40% provided by Palmer et al. (2006), who used a combination of six years of satellite observations (40 km × 320 km) and model simulations (2º × 2.5º) to estimate the interannual variability of isoprene emissions in the southeastern United States. Presumably, interannual variation in biogenic emissions on a regional scale is larger than that on a global scale, where estimates are effectively smoothed by averaging across all land.

Less attention has been paid to providing high-resolution, regional-scale assessments of biogenic emissions variability. The effects of BVOCs and their reaction products on cloud formation and atmospheric radiative transfer make them an important constituent to represent within future generations of regional climate models. Regional-scale climate impact assessments are of value to policymakers (e.g., air-quality managers, urban planners), who design regulations and resource-management plans for localities and regions.

Regardless of the size of the modeling domain, most efforts to model BVOC emissions employ the empirical algorithm of Guenther et al. (1995), which simulates emission of isoprene (the most dominant BVOC species), monoterpene, other volatile organic compounds (VOCs), other reactive VOCs, and carbon monoxide (which it treats as if it were emitted from vegetation). The 1995 Guenther algorithm represents BVOC flux as a function only of plant species, PAR, canopy temperature, and leaf biomass density:

$$F_{BVOC} = \varepsilon \gamma_{PAR} \gamma_T D \tag{4.1}$$

where F_{BVOC} is the land surface–to–atmosphere flux of BVOCs (units: μg C m^{-2} h^{-1}); ε is a vegetation-type–specific emission capacity (units: μg C $gdlm^{-1}$ h^{-1}, where gdlm is g dry leaf matter); γ_{PAR} is a dimensionless scalar that is a nonlinear function of PAR reaching the canopy surface (for non-isoprene BVOCs, $\gamma_{PAR} = 1$); γ_T is a dimensionless scalar that adjusts BVOC flux in response to changes in canopy temperature; and D is the leaf biomass density (units: gdlm m^{-2} of ground covered by the vegetation type). More recent modifications to the algorithm consider seasonal variation in plant emitting capacities and explicitly account for variation in soil moisture as a source of variability in BVOC emission rates (Guenther et al., 2006). For the research presented here, we assume that the Guenther et al. (1995) algorithm captures the key environmental drivers of emission variability.

Even if the Guenther et al. (1995) algorithm is assumed to perfectly represent the emissions processes, realistic simulation of BVOC flux also requires accurate representation of each of the three component sources of variation in the Guenther et al. algorithm: (1) short-term environmental variation (i.e., PAR reaching the leaf surface and canopy temperature); (2) seasonal and interannual changes in leaf-biomass density (phenology), which vary with changing environmental conditions (e.g., spring soil temperatures, plant-available soil moisture in the early summer); and (3) the species composition of a landscape.

Previous model-derived estimates of BVOC flux variability lay the foundation for future work (Levis et al., 2003; Naik et al., 2004; Tao and Jain, 2005), but the model frameworks used by previous researchers do not realistically represent all three component sources of variation listed above. Naik et al. (2004) and Tao and Jain (2005)

used monthly mean climate data to drive their models. Because biogenic emissions are a highly nonlinear function of PAR and canopy temperature, use of monthly mean climate data to drive a model likely underestimates total emissions and the variability of emissions. Naik et al. (2004) and Levis et al. (2003) used dynamic global vegetation models (DGVMs) to represent interannual phenological variation. The rate at which most species of trees emit biogenic emissions is one to two orders of magnitude higher than the rate at which grasses emit BVOCs. Accurate specification of the tree-to-grass ratio on the model domain landscape is consequently of critical importance when modeling biogenic emissions, but the skill of DGVMs to reproduce vegetation type composition at high resolutions has not been demonstrated. For example, the DGVM used by Levis et al. overestimates the ratio of grass to trees, most likely because of a dry bias in the coupled land-surface model (LSM) soil profile (Bonan and Levis, 2006). Furthermore, DGVMs may not accurately represent vegetation dynamics, phenology, or carbon and water fluxes at high model resolutions (Kucharik et al., 2006).

We sought to employ the most realistic currently available representations of the processes to which biogenic emissions estimates are most sensitive on time scales of years to decades (i.e., meteorological variation, seasonal and interannual variation in leaf-biomass density, vegetation composition of the model landscape). We employed a process-based, short-term dynamic phenology module within the National Center for Atmospheric Research's Community Land Model (CLM) (Bonan et al., 2002a; Bonan et al., 2002b; Levis et al., 2004; Oleson et al., 2004), a current-generation land-surface model (LSM) that contains a biogenic emissions module (Levis et al., 2003). LSMs allow biomass density, PAR, and canopy temperature to vary within a physically consistent

framework and can be used for both retrospective analysis and predictive research. Gulden and Yang (2006) showed that when equipped with species-based regional emissions capacities for LSM land-cover types, regional LSMs adequately reproduce the spatial distribution and magnitude of BVOC flux when compared to species-based emissions modules. Our framework realistically represents meteorological forcing, updating model state variables with new meteorological input every hour. After preliminary model calibration, the simulated phenology is consistent with observed phenological variation. The species distribution of the model landscape is also realistic: to initialize the model, we used land-cover data and biogenic emissions factors that were derived from the same high-resolution, ground-survey–derived, species-based land-cover dataset (Wiedinmyer et al., 2001; Gulden and Yang, 2006). Although climate-driven vegetation composition change over many centuries can result in a several-fold change in biogenic emissions (Lathière et al., 2005), we assume that over the course of years to decades, the relative plant-type distribution of a landscape can be considered effectively constant.

The augmented CLM (as described above) was used to address the following questions: (1) On a regional scale, how much do biogenic emissions for a given month vary from year to year? (2) Approximately what portion of this variation is due to direct climate variation (e.g., changes in PAR reaching the canopy, changes in leaf-surface temperature)? (3) Approximately what portion of this variation can be attributed to interannual changes in the amount of leaf biomass resulting from short-term variation in environmental conditions (e.g., more rain, warmer spring temperatures)? We present a model framework that can be used to represent the response of biogenic emissions to

changing climate conditions on a regional scale. Our method is demonstrated for eastern Texas, but it can be readily applied elsewhere.

4.3. METHODS AND MODELS

We coupled a short-term dynamic phenology model (Dickinson et al. 1998) to CLM and used the biogenic emissions module added by Levis et al. (2003) (see Equation 4.1; Guenther et al., 1995). Estimates of total BVOCs emitted globally using CLM with Levis et al.'s BVOC emissions module are consistent with numerous other estimates (e.g., Guenther et al., 1995; Wang and Shallcross, 2000).

We simulated biogenic emissions on a 0.1º grid over eastern Texas (28ºN to 33.5ºN; –99.5E to –94.5E) (Figure 4.1) using a 10-km, species-based land-cover dataset (Gulden and Yang, 2006). Vegetation composition varies considerably across the domain: a mixed broadleaf and needleleaf evergreen forest covers far eastern Texas; a mosaic of woody savannas, grassland, and cropland covers the rest of the domain. There is a strong west–east gradient in mean annual precipitation (Figure 4.2), which is reflected in the prominent west–east gradient in biomass density (Figure 4.3).

4.3.1. The Community Land Model

CLM is representative of a current-generation LSM that is used in climate and weather research. It simulates flows of mass, energy, and momentum between different reservoirs of the land surface. CLM represents canopy radiative transfer (Dickinson, 1983; Sellers, 1985; Bonan, 1996), photosynthesis and stomatal conductance (Farquhar et al., 1980; Collatz et al., 1991; Collatz et al., 1992; Dougherty et al., 1994), and transpiration (Oleson et al., 2004). It can be driven as part of a fully coupled climate

system model or "offline" using preexisting, high-resolution meteorological forcing data. Model time steps are short (minutes to hours).

4.3.2 Dynamic phenology module

To enable simulation of interannual variation in leaf biomass density without having model results depend on CLM-DGVM's spuriously high ratio of grass to trees (Bonan and Levis, 2006), we replaced CLM's DGVM with a dynamic phenology module (Dickinson et al., 1998). Unlike CLM-DGVM, the augmented LSM (CLM-DP) keeps the plant functional type (PFT) fraction of a grid cell equal to that specified at model initialization, but it allows leaf biomass density to vary as a function of soil moisture, soil temperature, canopy temperature, and vegetation type.

The dynamic phenology module allocates carbon assimilated during photosynthesis to leaves, roots, and stems; the fraction of photosynthate allocated to each reservoir is a function of the existing LAI. LAI is a linear function of leaf biomass density of the landscape; we used PFT-specific leaf biomass densities derived in concert with the initialization dataset (Gulden and Yang, 2006). The model tracks growth and maintenance respiration, represents slow-turnover and fast-turnover carbon reservoirs, and simulates vegetation response to cold stress and drought stress. Because the dynamic phenology module represents LAI as a nonlinear function of multiple environmental variables, the relationship between climate and LAI is not straightforward: in general, wetter soil, higher rates of photosynthesis, warm soil, and temperate canopy air result in increased LAI. Dickinson et al. (1998) provides a detailed description of the model.

Whether the Dickinson et al. or an equivalent dynamic phenology model is adequate to realistically represent changes in leaf biomass is not well established (Kim and Wang, 2005). Our review of the ecological literature leads us to believe that there is not yet firm consensus in the ecological research community regarding the quantification of factors that control short-term variation in biomass (e.g., Grier and Running, 1977; Gholz, 1982; Leuschner et al., 2006). The Dickinson et al. (1998) scheme was based on the best-available understanding of carbon allocation between vegetation reservoirs and is therefore a reasonable choice for representing short-term phenological variation. Dickinson and colleagues' model does not represent long-term competition between vegetation types and is not susceptible to model-produced errors in vegetation-type distribution, which pose a potentially significant error source for estimates of biogenic emissions.

4.3.3. Modifications to model processes and parameters

So that timing and shape of the seasonal cycle of the dynamically simulated LAI would better match those of the satellite-observed LAI, we changed the shape of the function allocating assimilated carbon to leaves (Figure 4.4) and re-coded the model so that leaves receive no assimilated carbon outside of the model growing season. Our selection of carbon-allocation function was guided by the function's influence on the shape of the seasonal cycle of the modeled LAI. Following CLM-DGVM, we treated the stem area index (SAI) as a constant fraction of LAI. SAI represents non-woody, non-leaf biomass above ground; as SAI increases, the shaded fraction of the canopy also increases.

To ensure that modeled seasonal cycle and magnitude of seasonal variation in LAI were consistent with observations, we manually calibrated the parameters of CLM-DP. We used two satellite-derived monthly time series as visual constraints (Figure 4.5): (1) the domain-averaged LAI for 1982–2000, which we calculated from 0.5º×0.5º Advanced Very High Resolution Radiometer (AVHRR) data and (2) the domain-averaged climatological LAI (i.e., the mean annual cycle averaged from data spanning 2001–2003) obtained from Moderate Resolution Imaging Spectroradiometer (MODIS) and AVHRR data (Lawrence and Chase, 2007).

The MODIS climatological LAI provides no information about interannual variability. AVHRR LAI data for 1982–2000 provided us with a lengthy, continuously varying time series. However, AVHRR-derived LAI data likely overestimates the magnitude of the mean LAI in eastern Texas (Tian et al., 2004; Lawrence and Chase, 2007) and underestimates interannual variation in LAI because it is based on the Normalized Difference Vegetation Index, which saturates quickly, especially in forested regions like eastern Texas (e.g., Wang et al., 2005). We used the MODIS climatological LAI time series as the target for the mean magnitude and amplitude of the seasonal cycle; AVHRR LAI data informed our understanding of the variability of LAI in the model domain.

We adjusted parameters until the seasonal cycle of LAI and the timing of leaf-on and leaf-off had roughly the same shape as the satellite-derived data. We uniformly scaled the specific leaf area values for PFTs calculated by Gulden and Yang (2006) by 1.5. For each PFT, we increased the minimum LAI from 0.05 m^2 m^{-2} to one half of the domain-wide minimum values reported in the MODIS-derived land-cover dataset

(Lawrence and Chase, 2007). For woody vegetation, we set SAI equal to 0.25×LAI; for grasses, we set SAI equal to 0.05×LAI. CLM uses SAI to calculate the shaded fraction of the canopy; the dependence of the rate of photosynthesis on the shaded fraction makes the modeled LAI particularly sensitive to the SAI fraction. Other slight adjustments to parameters were used if they improved the visual match of the datasets to the satellite-derived data; however, model output was insensitive to most parameters.

Figure 4.5 shows observed, uncalibrated, and calibrated domain-average LAI. The average absolute percent departure from the mean LAI (see Appendix for definition) provides a measure of variability that is not directly tied to the magnitude of LAI. Table 4.1 provides the average and maximum absolute percent deviation from the monthly mean domain-average LAI. Calibration of the dynamic phenology module decreased the average absolute percent deviation from the monthly mean LAI from 15.2% to 12.9%, which made the variability of the simulated LAI more consistent with that of the AVHRR dataset. Manual calibration improved simulated LAI across the domain: the calibrated, quadrant-averaged LAI for each of the four quadrants of the domain (see Figure 4.1) is consistent with the quadrant-averaged satellite observations (data not shown). Anomalies of the calibrated LAI are consistent in both sign and magnitude with those of the satellite-derived data (results not shown).

LAI is lowest in hot, dry years and highest in wet, temperate years (Fig. 4.6). Year-to-year differences in precipitation appear more important in determining LAI variation than do year-to-year differences in temperature.

4.3.4. Model runs

We report the results of two model runs: (1) One run used the standard version of CLM (CLM-PRESC); for each model year, it prescribed the MODIS climatological LAI. (2) The second run used CLM-DP; it simulated LAI variation in response to changing meteorological conditions. For both runs, we specified the plant functional type (PFT) composition of the landscape using a 0.1º, land-cover dataset that Gulden and Yang (2006) derived from a 1-km species-based dataset that was compiled from extensive ground surveys (Wiedinmyer et al., 2001). Gulden and Yang (2006) used the species-based dataset to compute species-based, region-specific biogenic emissions factors for each PFT (ε in Equation 4.1), which we used in both runs.

We bilinearly interpolated North American Regional Reanalysis (NARR) data (Mesinger et al., 2006) to create 0.1º meteorological forcing. We simulated the period 1979–2004, using a one-hour model time step. Simulation years 1979–1981 served as the spin-up period. We analyzed monthly mean model output from model years 1982–2004. The meteorological forcing was the same for both runs.

4.4. RESULTS AND DISCUSSION

A time series of simulated domain-mean BVOC flux is shown in Figure 4.7. Because the uncertainty in the magnitude of BVOC emissions probably exceeds a full order of magnitude (e.g., Guenther, 1997; Simpson et al., 1999; Smiatek and Bogacki, 2005), we report simulated variability using the average absolute percent departure from the monthly mean, a measure of variability that is not directly tied to the magnitude of the simulated flux (see definition in Appendix). To provide information about absolute flux,

we also report the standard deviation and its mean-normalized counterpart, the coefficient of variation (CV).

The domain-mean average absolute departure from the monthly mean BVOC flux was 22.4% when we used CLM-DP and 11.7% when we used CLM-PRESC (Figure 4.8 and Table 4.2); the domain-maximum average absolute departure from the monthly mean BVOC flux increased from 70.6% (using CLM-PRESC) to 137.7% (using CLM-DP) (Table 4.2). When only the BVOC fluxes from June, July, and August (JJA) were considered, the domain-mean (domain-maximum) average absolute departure from the monthly mean BVOC flux fell to 15.7% (35.3%) when we employed CLM-DP and to 7.0% (23.0%) when we used CLM-PRESC. A similar decrease was observed in each of the four quadrants (Figure 4.1) of the domain (results not shown).

We present detailed results for JJA because biogenic emissions peak in the summer months, concurrent with the summer ozone season in Texas. This is the time period when the biogenic emissions may play the most important role in the regional atmospheric chemistry. It is important to note that some BVOC-related climatic processes may be most sensitive to relative variability during other seasons; however, because such processes are poorly understood, we focus on the period of the year when absolute BVOC flux is greatest.

The variation in leaf-biomass density is a significant source of year-to-year variation in JJA mean BVOC flux: the domain-average standard deviation of the seasonal (JJA) mean BVOC flux for the period 1982–2004 is 123 μg C m^{-2} h^{-1} when phenology is prescribed and 401 μg C m^{-2} h^{-1} when phenology is allowed to vary with environmental conditions. Use of dynamic phenology results in threefold increase in the domain-average

CV of JJA mean BVOC flux: the CV is 0.183 when CLM-DP is used and 0.0619 when CLM-PRESC is used.

Model results are qualitatively consistent with expectations: biogenic emissions peak in the evergreen–deciduous mixed forests of eastern Texas, where biomass density is highest and where the dominant vegetation species have high emission capacities (Fig. 4.9.A). Figure 4.9.B1 and Figure 4.9.B2 show the spatial distribution of the average absolute departure from the JJA mean flux; Figure 4.9.C1 and Figure 9.C2 show the spatial distribution of the standard deviation of the JJA mean flux. Grid cells where variability is highest are often grid cells in which there are relatively few trees (e.g., in the northwest corner of the domain) (Figure 4.3) and where annual precipitation is relatively low but also highly variable between years (Figure 4.2). It logically follows that interannual variation in BVOC flux is highest in areas where the absolute BVOC flux is relatively low: as the JJA mean BVOC flux increases, the grid-cell's CV of JJA mean BVOC flux becomes less sensitive to interannual variation in LAI (Figure 4.10).

Modeled isoprene emissions are more sensitive than modeled monoterpene emissions to use of dynamic phenology (Table 4.2). Modeled isoprene flux is a nonlinear function of both PAR and canopy temperature (both of which are altered by changing LAI), whereas monoterpene flux responds only to changes in canopy temperature.

Even though BVOC flux is a linear function of biomass density (Equation 4.1), there is not a one-to-one correspondence between the CV of the JJA mean LAI and the CV of the JJA mean BVOC flux (Figure 4.11). LAI nonlinearly alters both PAR and canopy temperature. Increasing LAI increases light attenuation through the canopy, which decreases the amount of PAR reaching the lower levels of the canopy and

consequently decreases the PAR-dependent emission scaling factor, γ_{PAR}. Using CLM-DP instead of CLM-PRESC increases the standard deviation of canopy temperature (Figure 4.12), which augments variability of the temperature-dependence emission scaling factor, γ_T. The relationship between canopy temperature and variation in LAI is not straightforward: increasing leaf area increases the latent cooling capacity of the canopy; however, increasing leaf area also increases the canopy radiative absorption capacity. Because the simulated canopy temperature during June, July, and August was, on average, 0.19 K cooler when dynamic phenology was employed than when phenology was prescribed, we surmise that the increased latent cooling capacity of the canopy is the dominant source of increased variability in canopy temperature and, consequently, in biogenic emissions.

To quantitatively assess the relative contribution of short-term environmental variation (e.g., shifts of canopy temperature) and the relative contribution of interannual variation in LAI to year-to-year differences in the magnitude of BVOC flux, we would need to perform parameter-substitution experiments within simulations using a coupled land–atmosphere model that allows changes in albedo, transpiration, and surface roughness to feed back to affect regional meteorological conditions. However, we can approximate the relative importance of variation in temperature, radiation, and biomass to the interannual variability of BVOC flux by comparing the component sources of variability from CLM-PRESC to those from CLM-DP run. Figure 4.13 shows the domain-average percent departure from the monthly mean for isoprene and the corresponding domain-average percent departure from the monthly mean for the dimensionless emission-modulating factors γ_{PAR} and γ_T and for LAI. We see in Figure

4.13 that LAI is a source of variability that is at least of equal importance to PAR and temperature. When seeking accurate assessments of interannual variation in non-isoprene BVOC emissions, which (at least in the model) respond only to variation in temperature, simulating dynamic phenology is at least as important.

In conjunction with the findings of others (Levis et al., 2003; Naik et al., 2004; Tao and Jain, 2005), our results reinforce the assertion that total biogenic emissions vary significantly between years—by up to a factor of two. We demonstrate the importance of representing of interannual change in leaf-biomass density when modeling biogenic emissions, and we show the need for improved empirical quantification of interannual variation in leaf-biomass density. In Texas, year-to-year variation in precipitation is more important than temperature in controlling the interannual variation of leaf-biomass density: our results supply additional justification for studies that predict future changes to interannual variability of precipitation.

Even minor changes in the concentrations of BVOCs can shift the tropospheric ozone production of a region from being NO_x-limited to being VOC-limited, a chemical shift that changes the relative utility of air-pollution control policies (Wang et al., 2005). For the sake of simplicity, air-quality modeling studies often assume that the annual phenological cycle is constant between years and thereby likely underestimate interannual differences in biogenic emissions and air quality.

The preliminary calibration used here provides improved estimates of the magnitude and spatial distribution of year-to-year variation of biogenic emissions in eastern Texas. Such information is immediately useful to Texas air-quality managers and policymakers; however, automated multicriteria calibration (e.g., Bastidas et al., 1999) of

model parameters will further improve the realism of simulated emissions, will provide uncertainty estimates, and will improve the model's suitability for practical and policy applications.

CLM's biogenic emissions module depends on the premise that biogenic emissions are a linear function of leaf-biomass density. During our preliminary calibration, when we treated the satellite-derived LAI data as truth, we implicitly assumed that leaf biomass density is a linear function of satellite-observed spectral reflectance. The combination of these two assumptions implies that leaf pigmentation and BVOC flux are correlated, consistent with empirical studies (Lehning et al., 2001) but which may not be completely accurate.

The model framework used here neglects many sources of interannual variation in biogenic emissions. Geron et al. (2000) showed that emissions factors for white oak vary seasonally; Funk et al. (2003) found evidence for large variation in emission factors over the daily cycle. CLM does not represent the detrimental effect of ozone on plant growth (Ashmore, 2005), and it does not account for recent observations showing a decrease in the emission of some chemical species of BVOCs when ambient carbon dioxide concentrations increase (e.g., Rosenstiel et al., 2003). The BVOC emissions module in CLM does not represent the uptake of atmospheric isoprene from soil microorganisms (Cleveland and Yavitt, 1998), the rate of which appears to decrease as soil moisture decreases (Pegoraro et al., 2005). At least in some plant species, isoprene emissions increase under mild drought stress and then decrease when drought stress becomes severe (Pegoraro et al., 2005; Funk et al., 2005). In CLM-DP, LAI begins to decrease as soon as photosynthesis falls below the maintenance respiration rate; all else being equal, modeled

isoprene flux begins to decrease as soon as maintenance respiration demands more energy than is captured by photosynthesis. CLM-DP therefore likely underestimates BVOC flux in periods of mild water scarcity.

Why bother to augment an LSM with a dynamic phenology model when high-resolution satellite-derived observations of LAI are available? Assimilating high-resolution satellite-derived LAI variation is indeed a viable option for retrospective applications. The work of Guenther et al. (2006) provides a ready framework for such assimilation. However, such an approach is not amenable to investigation of how future environmental change will alter biomass density and biogenic emissions. Processed-based representations of phenological change also provide a foundation for future assessment of cause-and-effect relationships between BVOCs, leaf biomass density, and other components of the climate system; using prescribed, satellite-derived data would render such investigation impossible.

A dearth of observations limits realistic simulation of biogenic emissions. Region-specific biogenic emissions capacities for the coarse-resolution LSM land-cover classifications and any estimates in interannual variability of biogenic emissions are, of course, only as good as the species-based information from which they are derived. Although the number of species for which the scientific community has biogenic emissions rate estimates is growing, most plant species have no associated biogenic emissions data and data quantifying intraspecies variation in both type (Staudt et al., 2004) and quantity (e.g., Funk et al., 2005) of biogenic emissions is available for only a very few vegetation species.

4.5. SUMMARY AND CONCLUSIONS

CLM augmented with a dynamic phenology module, a species-derived land-cover dataset, and region-specific biogenic emissions factors were used to estimate the interannual variability of biogenic emissions in eastern Texas from 1982–2004.

Using standard CLM, the domain-mean (domain-maximum) average absolute departure from the monthly mean BVOC flux is 11.7% (70.6%); using CLM with dynamic phenology it is 22.4% (137.7%). The domain-mean (domain-maximum) average absolute departure from the monthly mean flux is lower during summer: it is 15.7% (35.3%) using dynamic phenology; it is 7.0% (23.0%) using prescribed phenology. When phenology is prescribed, the domain-average coefficient of variation (mean-normalized standard deviation) of the JJA mean BVOC flux is 0.0619. When phenology is allowed to vary with environmental conditions, the domain-average coefficient of variation increases threefold to 0.183.

Interannual variation in leaf-biomass density is a significant source of year-to-year variation in regional biogenic emissions that is at least as important as interannual variation in temperature or photosynthetic active radiation. Phenology-driven biogenic emission variability is greatest in regions with relatively low absolute emissions: as a grid cell's mean BVOC flux increases, the mean-normalized standard deviation of mean BVOC flux tends to decrease. Variability is highest in sub-humid, sparsely wooded regions where interannual variability of precipitation is relatively large.

Our model framework paves the way for future coupled climate model investigations of the cause-and-effect relationships between biogenic emissions, other

land-surface states and fluxes, and atmospheric processes. Our results are for Texas, but they likely can be extrapolated elsewhere.

4.6. ACKNOWLEDGEMENTS

We are particularly grateful to Peter Lawrence, who provided us with the satellite-derived land-cover dataset and who lent us many useful insights. We are also grateful for the constructive comments of three anonymous reviewers, David Allen, Mark Estes, Enrique Rosero, Xiaoyan Jiang, Xuemei Wang, Robert Dickinson, and Hua Su. This work was funded by the U.S. Environmental Protection Agency's Science to Achieve Results (STAR) Program Grant #RD83145201. Thanks to the Texas Advanced Computing Center for providing computing resources.

4.7. CHAPTER 4 APPENDIX

We define the absolute percent departure from the mean for a time series x as follows:

$\forall x_{y,m}$; $y = y_s, y_{s+1}, \ldots, y_e$; $m = 1, 2, \ldots, 12$

$$\overline{x_m} = \frac{\sum_{y=y_s}^{y_e} x_{y,m}}{(y_e - y_s + 1)} \tag{4.A.1}$$

$$d_{y,m} = \left| \frac{x_{y,m} - \overline{x_m}}{\overline{x_m}} \right| \tag{4.A.2}$$

$$\overline{d_{m_s - m_e}} = \frac{\sum_{y=y_s}^{y_e} \sum_{m=m_s}^{m_e} d_{y,m}}{(y_e - y_s + 1) \times (m_e - m_s + 1)} \tag{4.A.3}$$

$$\max(d_{m_s - m_e}) = \max(d_{y,m}),\ y = y_s, y_{s+1}, \ldots, y_e,\ m = m_s, m_{s+1}, \ldots, m_e \quad (4.A.4)$$

Where $x_{y,m}$ is a member of time series x. $x_{y,m}$ is the monthly mean value of the variable of interest for month m of year y (Equation 4.A1). The first and last years of time series x are y_s and y_e (1982 and 2004, in the results described in this paper). For a given month m in year y, that month's absolute departure from the monthly mean is $d_{y,m}$ (Equation 4.A2). The average absolute departure from the monthly mean for the time series x considering only the set of months that begins with m_s and ends with m_e is $\overline{d_{m_s - m_e}}$ (Equation 4.A3). The maximum absolute departure from the monthly mean for the time series considering only the set of months that begins with m_s and ends with m_e is $\max(d_{m_s - m_e})$ (Equation 4.A4). To compute the domain-mean (domain-maximum) average absolute departure from the mean, we first calculated the average absolute departure from the mean for each grid cell and then calculated the mean (maximum) of the values for all grid cells that compose the domain.

Table 4.1. Observed and model-simulated leaf area index (LAI): Average and maximum absolute percent departure from the domain-averaged monthly mean.*

		Whole domain	Quadrant 1	Quadrant 2	Quadrant 3	Quadrant 4
AVHRR†	Avg.	9.00%	6.10%	15.00%	18.00%	8.70%
	Max.	*30.70%*	*29.30%*	*68.10%*	*109.40%*	*30.40%*
Uncalibrated CLM-DP	Avg.	15.20%	9.30%	10.40%	23.30%	29.30%
	Max.	*49.60%*	*57.90%*	*49.40%*	*90.80%*	*90.20%*
Calibrated CLM-DP	Avg.	12.90%	9.70%	11.40%	18.40%	22.90%
	Max.	*59.30%*	*62.80%*	*61.30%*	*78.30%*	*59.90%*

* The average and maximum absolute percent departure from the monthly mean is defined in the Appendix.

† The AVHRR LAI series, which is on a 0.5º grid, extends only from 1982–2000; the model-simulated LAI series, which are simulated on a 0.1º grid, extend from 1982–2004.

Table 4.2. VOC flux simulated using prescribed phenology and dynamic phenology: Average and maximum absolute percent departure from monthly mean BVOC flux.*

		Whole domain	Quadrant 1	Quadrant 2	Quadrant 3	Quadrant 4
		Total† BVOC flux				
Prescribed phenology	Avg.	11.70%	1.32%	13.50%	12.60%	7.30%
	Max.	*70.60%*	*74.90%*	*105.60%*	*78.90%*	*26.40%*
Dynamic phenology	Avg.	22.40%	20.90%	19.40%	27.60%	21.70%
	Max.	*137.70%*	*225.10%*	*159.90%*	*129.10%*	*70.70%*
		Isoprene				
Prescribed phenology	Avg.	12.70%	14.80%	13.20%	13.20%	8.20%
	Max.	*89.30%*	*97.70%*	*139.40%*	*91.10%*	*29.60%*
Dynamic phenology	Avg.	25.50%	25.00%	24.10%	30.60%	22.40%
	Max.	*203.70%*	*331.40%*	*288.00%*	*150.20%*	*80.60%*
		Monoterpene				
Prescribed phenology	Avg.	9.90%	10.50%	12.40%	11.10%	5.30%
	Max.	*42.30%*	*49.90%*	*67.90%*	*54.50%*	*21.10%*
Dynamic phenology	Avg.	19.20%	15.90%	16.00%	25.20%	19.90%
	Max.	*79.00%*	*120.20%*	*80.70%*	*91.60%*	*50.90%*

* See Appendix for explanation of average and maximum absolute departure from the monthly mean.

† Total VOC flux is the sum of the fluxes of isoprene, monoterpene, other volatile organic compounds, other reactive volatile organic compounds, and carbon monoxide (which CLM treats as if it were emitted by vegetation).

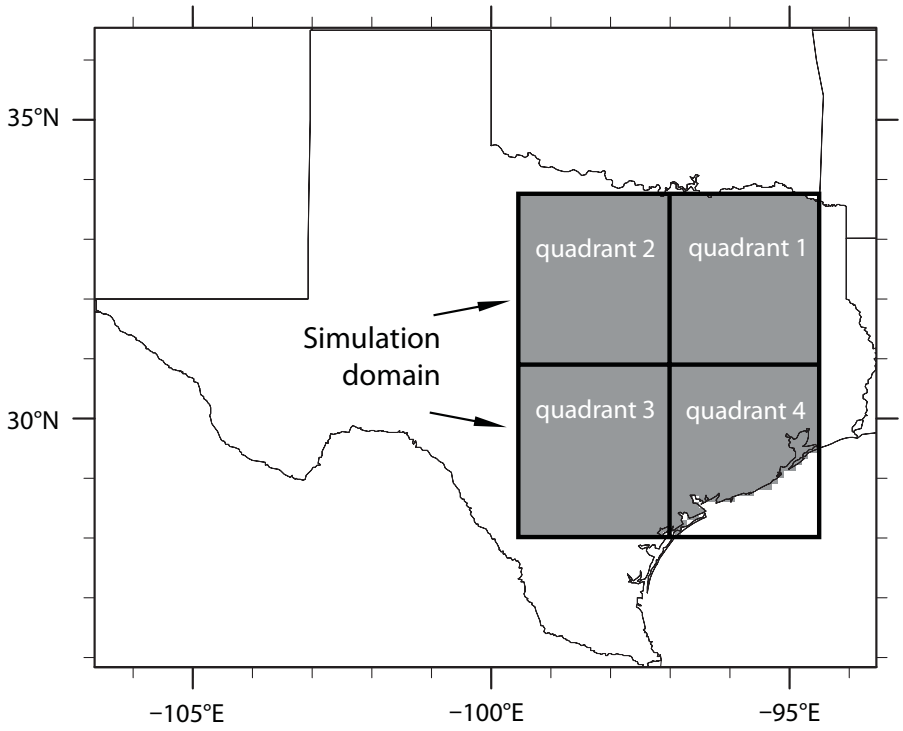

Figure 4.1. Model domain.

Quadrant-mean LAI values were used as an additional visual constraint when calibrating the dynamic phenology module within CLM (see description of calibration in text).

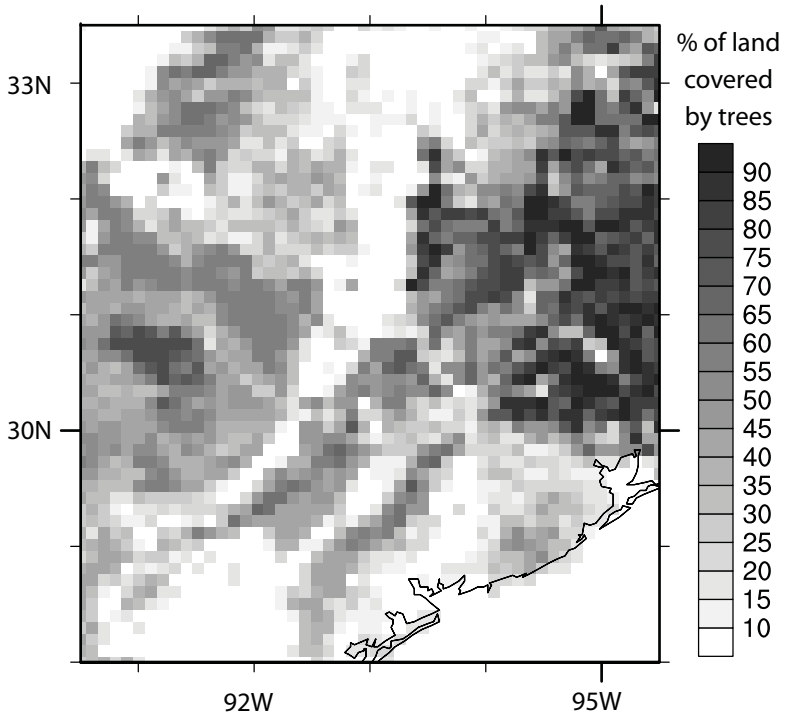

Figure 4.2. Annual rainfall characteristics of meteorological forcing used to drive model.

Data are bilinearly interpolated to 0.1º from North American Regional Reanalysis (NARR) data (Mesinger et al., 2006) for years 1982–2004. Note the dry bias of the NARR data along coastal Texas; resulting LAI values simulated with the dynamic phenology module are correspondingly spuriously small.

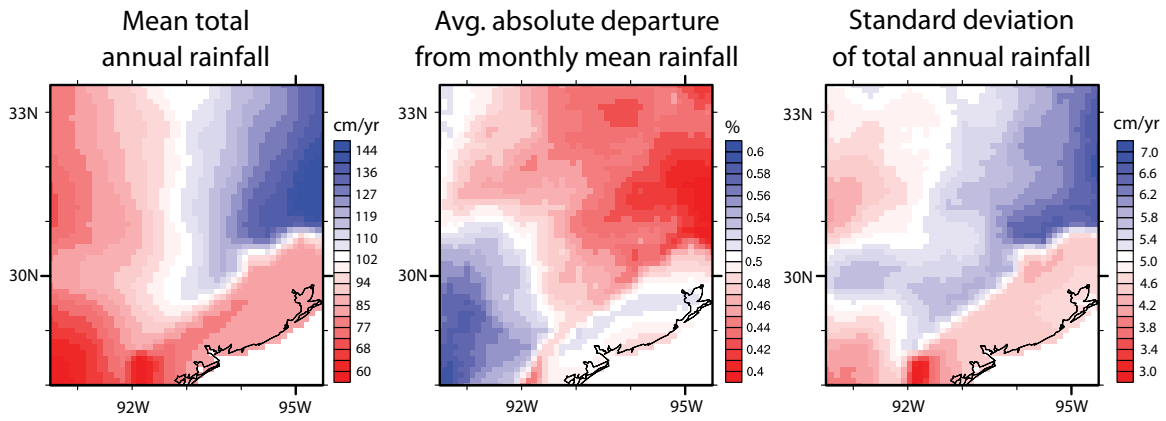

Figure 4.3. Percent of land area covered by trees.

Data shown were derived from the ground-survey–derived, vegetation-species–based land-cover dataset of Wiedinmyer et al. (2001) by Gulden and Yang (2006).

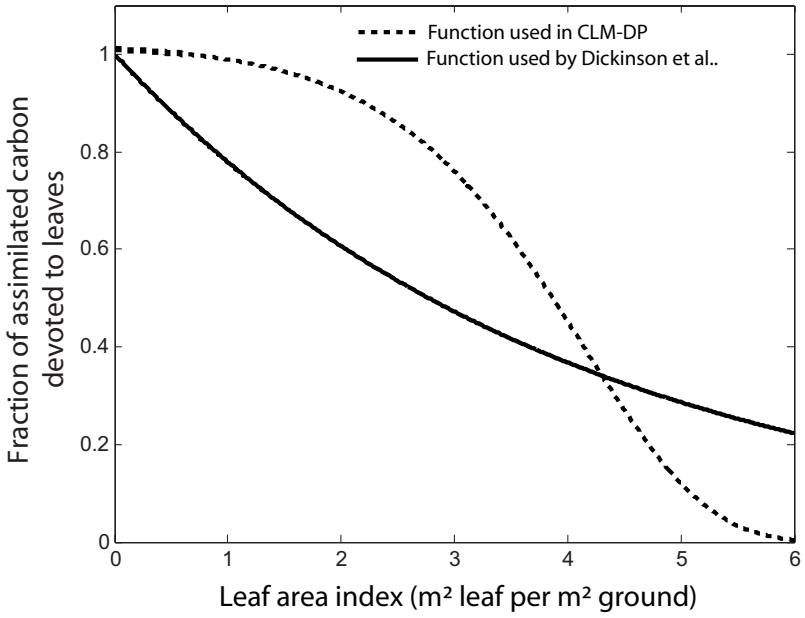

Figure 4.4. Fraction of assimilated carbon devoted to leaves as a function of leaf area index.

Comparison of the original function used by Dickinson et al. (1998) with that used in CLM-DP.

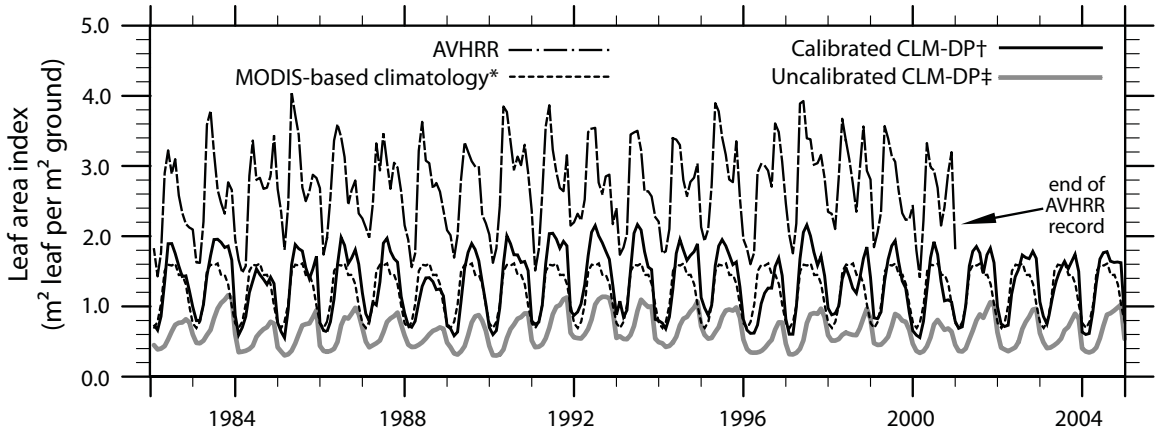

Figure 4.5. Comparison observed and simulated domain-average monthly leaf area index (LAI) for 1982–2004.

* Note that the MODIS-based climatology varies seasonally but not interannually.
† The "Calibrated CLM-DP" series is the LAI simulated by CLM-DP after manual calibration to the satellite-derived LAI and our modifications to the code (both of which are described in text)
‡The "Uncalibrated CLM-DP" series is the LAI simulated by CLM-DP without any modifications to the parameterizations or the parameters.

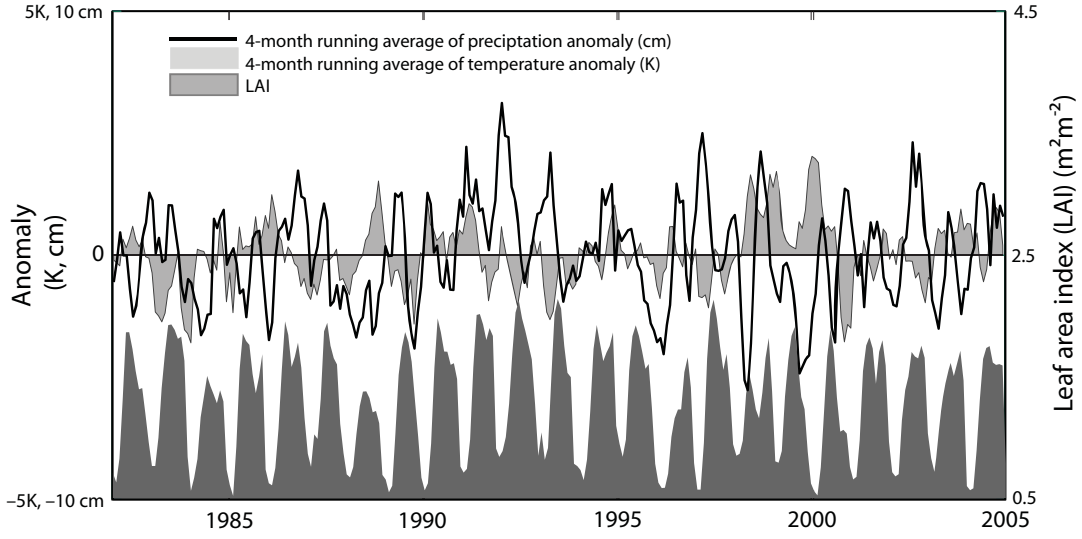

Figure 4.6. Domain-averaged temperature anomaly, precipitation anomaly, and leaf area index (LAI).

Temperature and precipitation anomalies were derived from the North American Regional Reanalysis (NARR) data (Mesinger et al., 2006) used to drive the offline models.

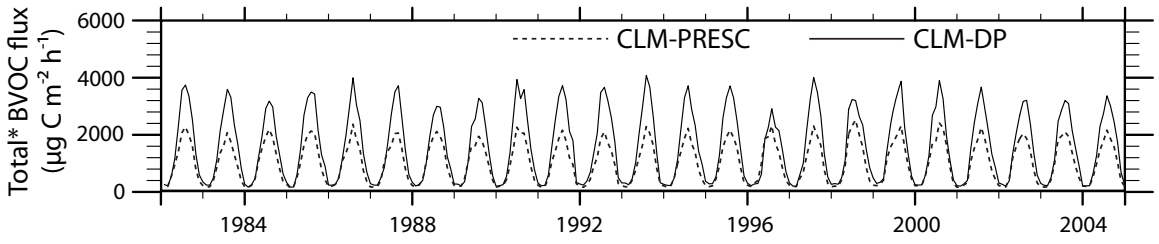

Figure 4.7. Time series of domain-averaged BVOC flux†.

*BVOC flux is the sum of the fluxes of isoprene, monoterpene, other reactive volatile organic compounds, other volatile organic compounds, and carbon monoxide (which CLM treats as if it were emitted by vegetation).
†Although both the CLM-PRESC and CLM-DP runs used the same ground-survey–based PFT distribution, CLM-PRESC relied on PFT-specific LAI values obtained from the MODIS-derived land-cover dataset (Lawrence and Chase, 2007). The ground-survey–based dataset has, on average, a higher area fraction of bare soil than the Lawrence and Chase dataset; consequently, CLM-PRESC mean LAI is smaller than the MODIS climatological observations. The magnitude of the BVOC flux simulated using CLM-PRESC is correspondingly lower.

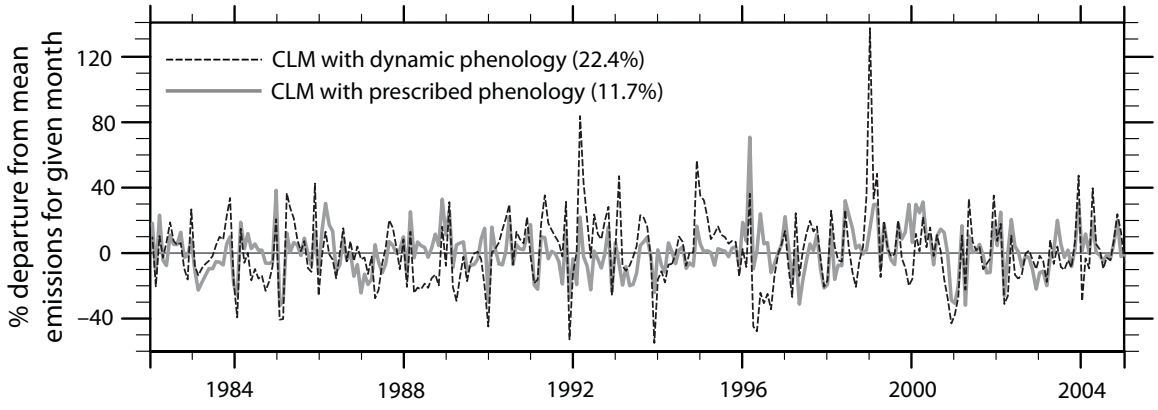

Figure 4.8. Percent departure from the monthly mean BVOC emissions.

Average absolute percent departure from the monthly mean is in parentheses. BVOC flux is the sum of the fluxes of isoprene, monoterpene, other volatile organic compounds, other reactive volatile organic compounds, and carbon monoxide (which CLM treats as if it were emitted by vegetation). See Appendix for definition of average absolute percent departure from the monthly mean.

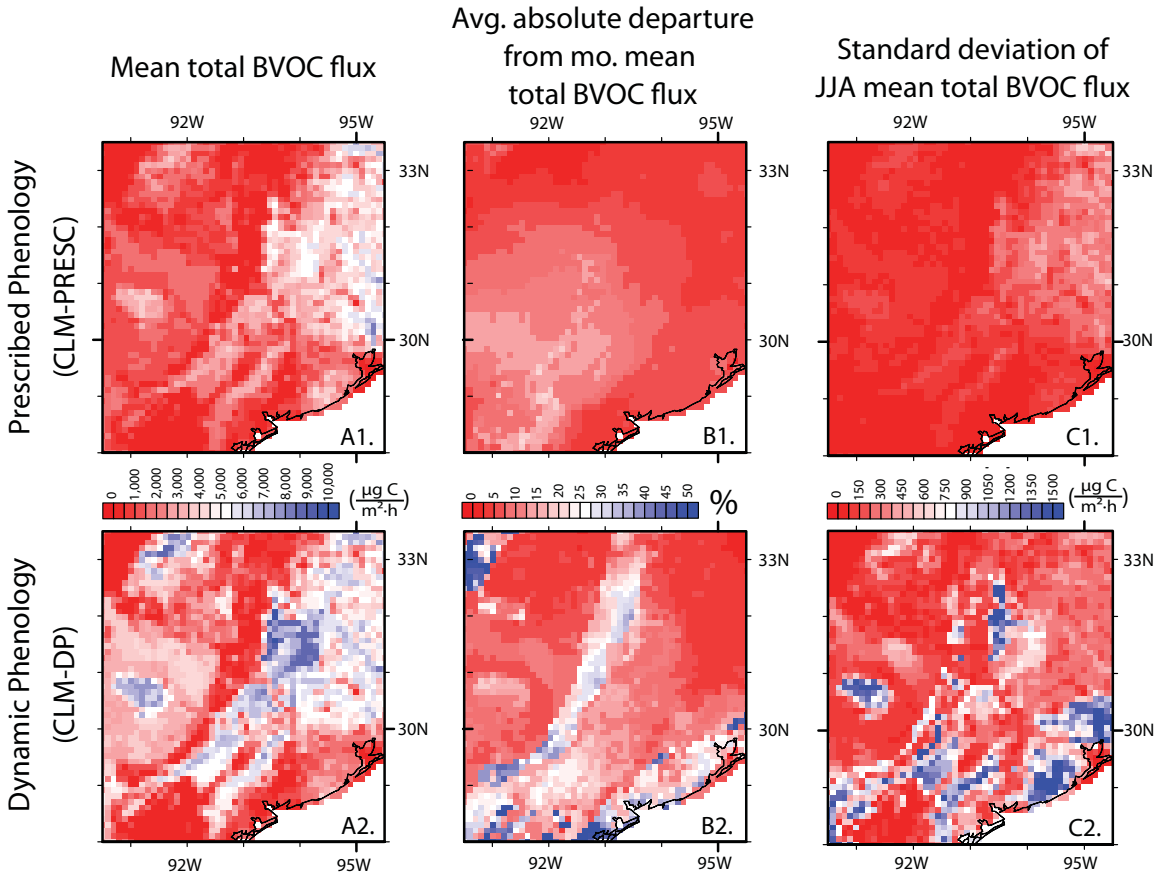

Figure 4.9. Summer BVOC flux (1982–2004 mean, average departure from the monthly mean, and standard deviation).

The standard deviation quantifies the interannual variation in the mean seasonal (mean of June, July, and August emissions) average BVOC flux. BVOC flux is the sum of the fluxes of isoprene, monoterpene, other reactive volatile organic compounds, other volatile organic compounds, and carbon monoxide. The Appendix explains the calculation of the average absolute departure from the monthly mean.

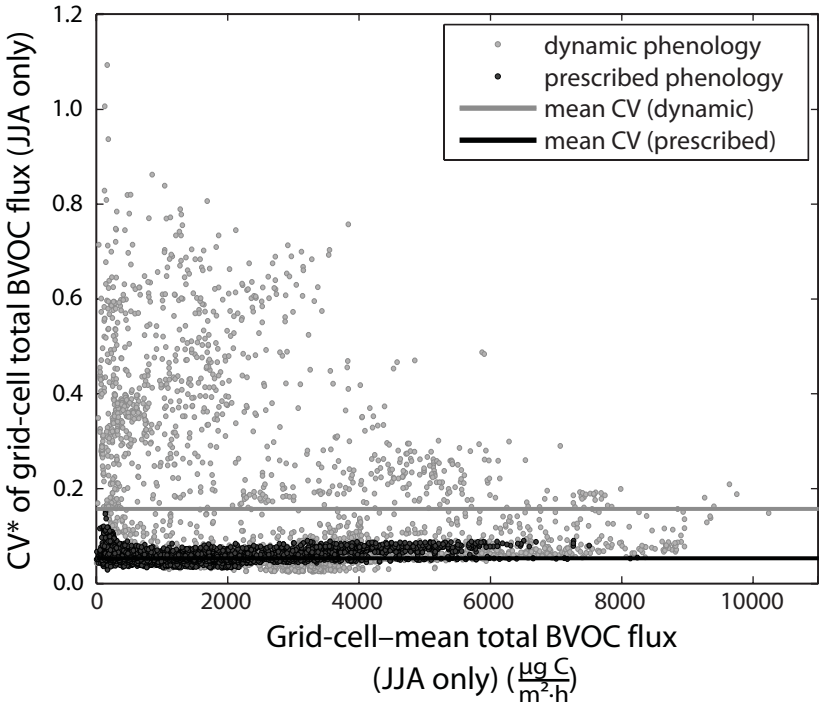

Figure 4.10. Relation of each grid cell's mean BVOC flux with its coefficient of variation (CV) for 1982–2004.

Only BVOC fluxes from June, July, and August are considered in the calculation of mean and CV. *CV is the standard deviation divided by the mean.

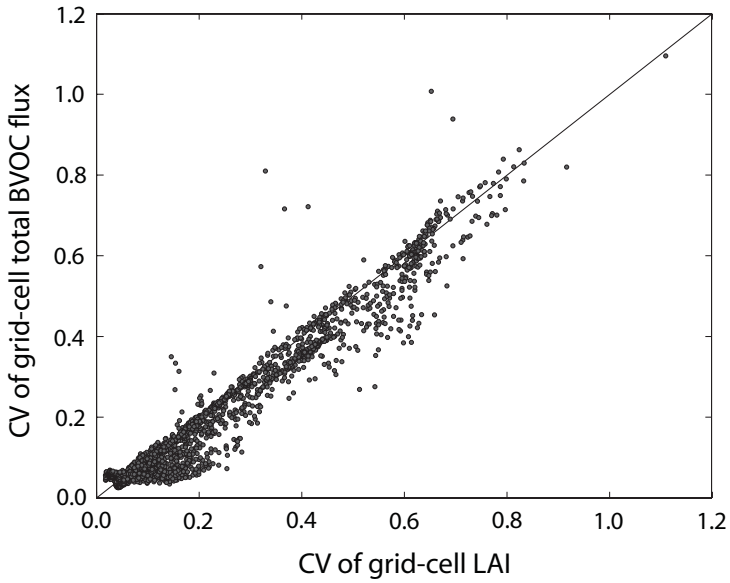

Figure 4.11. Relation between the coefficient of variation (CV) of leaf area index (LAI) and the CV of BVOC flux.

Results are from the simulation that used CLM-DP.

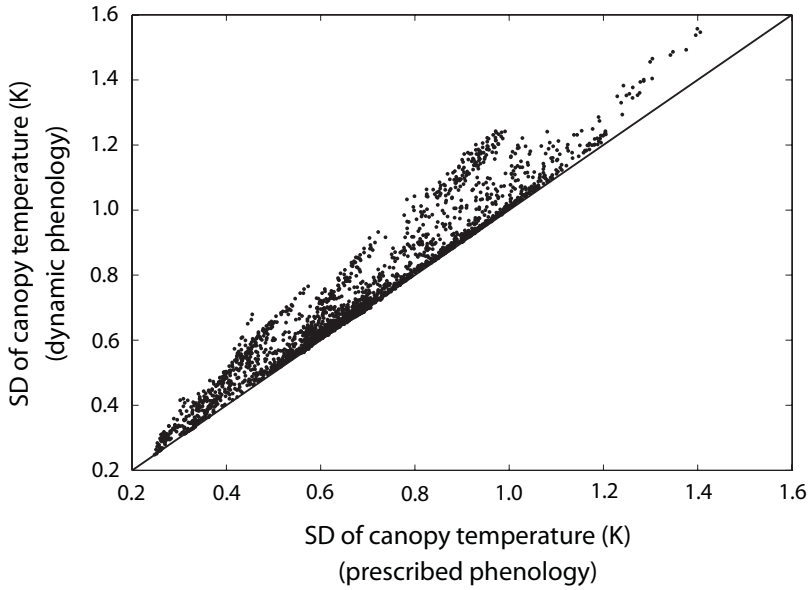

Figure 4.12. Relation between the standard deviation (SD) of canopy temperature in the CLM-DP run the CLM-PRESC run.

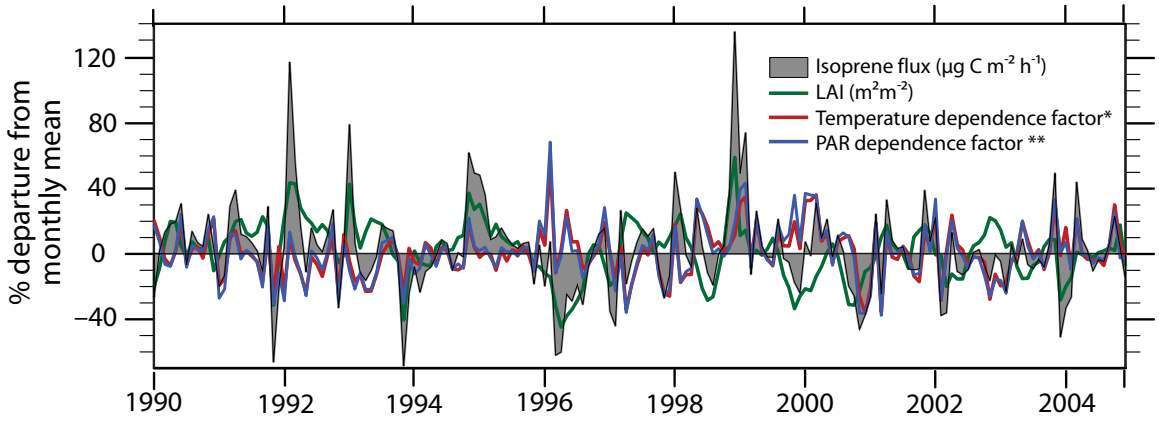

Figure 4.13. Components of interannual variation in domain-average isoprene emissions (1990–2004).

Time series shown are the domain-average % departure from the monthly mean. *The temperature dependence factor is γ_T in Equation 4.1. ** The PAR dependence factor is γ_{PAR} in Equation 4.1. The variability in the time period shown is representative of the variability for the entire simulation period; only years 1990–2004 are shown for ease of viewing.

Chapter 5: Improving the hydrology of land surface models: *Is an explicit aquifer model better than a deeper soil profile?* [4]

5.1. ABSTRACT

Monte Carlo analysis is used to show that explicit representation of an aquifer within a land-surface model (LSM) decreases the dependence of model performance on accurate selection of subsurface hydrologic parameters. Within the National Center for Atmospheric Research Community Land Model (CLM) we evaluate three parameterizations of vertical water flow: (1) a shallow soil profile that is characteristic of standard LSMs; (2) an extended soil profile that allows for greater variation in terrestrial water storage; and (3) a lumped, unconfined aquifer model coupled to the shallow soil profile.

North American Land Data Assimilation System meteorological forcing data (1997–2005) drive the models as a single column representing Illinois, USA. The three versions of CLM are each run 22,500 times using a random sample of the parameter space for soil texture and key hydrologic parameters. Other parameters remain constant. Observation-based monthly changes in state-averaged terrestrial water storage (dTWS) are used to evaluate the model simulations. After single-criteria parameter exploration, the schemes are equivalently adept at simulating dTWS. However, explicit representation

[4] Substantial portions of this chapter were originally published in Gulden, L.E., E. Rosero, Z. Yang, M. Rodell, C. S. Jackson, G. Niu, P. J.-F. Yeh, and J. Famiglietti (2007b), Improving land-surface model hydrology: Is an explicit aquifer model better than a deeper soil profile? *Geophys. Res. Lett.*, 34, L09402, doi:10.1029/2007GL029804. The References section contains full citations for all articles referenced here.

of groundwater considerably decreases the sensitivity of modeled dTWS to errant parameter choices. We show that approximate knowledge of parameter values is not sufficient to guarantee realistic model performance: because interaction among parameters is significant, they must be prescribed as a congruent set.

5.2. INTRODUCTION

With the growing recognition of groundwater–atmosphere interaction as a potentially significant influence on spatial and temporal climate variability, researchers in the field of terrestrial hydrometeorology have focused increasing attention on improving the process representations of subsurface hydrology within land-surface models (LSMs). Existing process representations fall within three broad classes: (1) multi-layered, relatively shallow soil columns in which groundwater storage is implicitly represented because the model conserves mass [e.g., *Oleson et al.,* 2004]; (2) many-layered, deep soil columns whose lower boundaries are beneath the climatological depth to the water table [*Koster et al.,* 2000; *Maxwell and Miller,* 2005]; (3) multi-layered soil columns coupled to lumped, unconfined aquifer models [*York et al.,* 2002; *Liang et al.,* 2003; *Yeh and Eltahir,* 2005; *Fan et al.,* 2007; *Niu et al.,* 2007].

Which of these methods best represents subsurface hydrology at a monthly time scale? We address this question for three different levels of parameter uncertainty: (1) when an optimal set of subsurface hydrologic parameters (e.g., percent sand, porosity, and specific yield) can be inferred from observations (the "ideal" case); (2) when no

information about effective parameters can be obtained (the "worst" case); and (3) when only ranges for parameter values are known (the "real life" case).

To ensure a fair comparison between methods, we isolate process representation as the primary source of uncertainty in model predictions. To limit input-data uncertainty, we employ the same meteorological forcing data and land-surface data for all runs. To explore the impact of parameter uncertainty, we use a Monte Carlo approach. Unlike calibration studies, the underlying goal of this work is not to identify the optimal parameter set; instead our primary goal is to evaluate and compare the added value of process representations.

Three questions frame our analysis: (1) When given a surrogate optimal parameter set, which of the ways to represent subsurface hydrology results in the most realistic simulation of monthly change in terrestrial water storage? (2) When no reliable information regarding effective subsurface hydrologic parameters exists, which process representation most consistently gives the best performance? (3) Does knowledge of approximate values for hydrologic parameters guarantee reasonably accurate simulation of monthly change in terrestrial water storage? Our results will inform LSM model development; more important, they characterize the level of confidence that can be placed in LSM-generated hydrologic predictions, especially when observations are scarce.

5.3. METHODS

We employ the National Center for Atmospheric Research's Community Land Model (CLM) [*Bonan et al.,* 2002; *Oleson et al.,* 2004; *Niu et al.,* 2005] as the host

model in which to test three methods for representing vertical water flow within the LSM soil column. The versions of CLM calculate surface and subsurface runoff (i.e., baseflow) as a function of topographic characteristics [*Niu et al.,* 2005] and are identical except for the method that they use to represent vertical water transfer in the soil column.

The first version of CLM (hereafter "SSOIL") uses the standard 10-layer, relatively shallow 3.43-m soil profile with topography-based runoff parameterizations [*Niu et al.,* 2005]. Because it conserves mass, the model implicitly represents groundwater dynamics; however, the true depth to the water table often exceeds the depth of the model's lower boundary. The second model (hereafter "DEEP") is identical to SSOIL except that it uses a 30-layer, 11.2-m soil profile, thereby extending the depth of the model soil profile to encompass a wider range of groundwater fluctuations. The third version (hereafter "AQUIFER") couples a lumped unconfined aquifer model to the standard 10-layer soil profile [*Niu et al.,* 2007]; it allows two-directional vertical water transfer between the unsaturated zone and the aquifer down a hydraulic gradient.

We run each model version as a single column representing the state of Illinois, USA. Illinois covers ~146,000 km^2. Crops and grass dominate the landscape. The climate is temperate and continental, and the topographic relief is relatively low. (See *Changnon et al.,* 1988 and *Yeh et al.,* 1998 for detailed descriptions of regional climate and hydrogeology.)

Meteorological forcing and land-surface input data are the area-weighted arithmetic averages of high-resolution datasets over the state of Illinois. The forcing is provided by the North American Land Data Assimilation System [*Cosgrove et al.,* 2003]. A CLM-compatible land-cover dataset derived from Advanced Very High Resolution

Radiometer and Moderate Resolution Imaging Spectroradiometer data [*Lawrence and Chase,* 2007] provides vegetation type distributions, biomass densities, and soil colors.

A Monte Carlo approach allows us to extensively explore the range of model responses across parameter space. We run SSOIL, DEEP, and AQUIFER 22,500 times each. A unique set of subsurface hydrologic parameters is used for each run. We randomly sample uniform or semi-uniform distributions that span physically reasonable ranges of values for soil texture parameters and other hydrologic parameters (Table 5.1). Each Monte Carlo run is initialized with a spun-up dataset created by running the model three times through the period 1997–2005 using default parameters. To allow for additional spin-up, the first year of each run is omitted from the analysis.

We assess the accuracy of model output using the statewide-average change in total column terrestrial water storage (dTWS), which we constructed from soil moisture and groundwater observations obtained by the Illinois State Water Survey (ISWS) [*Hollinger and Isard,* 1994; *Robock et al.,* 2000] following the methods of *Rodell and Famiglietti* [2001]. dTWS is a suitable constraint because it integrates the hydrologic behavior of the landscape; it is directly observable everywhere on Earth using Gravity Recovery and Climate Experiment (GRACE) measurements [*Chen et al.,* 2006]; and it properly represents the land storage term of the coupled atmospheric-terrestrial water budget.

Looking only at data from 1998–2005, we score parameter sets with the following metric:

$$F = \mathrm{RMSE} \times (1 - r) \tag{5.1}$$

where RMSE is the root mean square error between modeled and observed dTWS:

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(o_i - m_i\right)^2} \qquad \begin{array}{l} n = \text{length of time series} \\ o_i = \text{observed dTWS at time } i \\ m_i = \text{modeled dTWS at time } i \end{array} \qquad (5.1a)$$

and r is the correlation coefficient, defined as:

$$r = \frac{\sum_{i=1}^{n}\left(o_i - \overline{o}\right)\left(m_i - \overline{m}\right)}{\left(\sum_{i=1}^{n}\left(o_i - \overline{o}\right)^2 \sum_{i=1}^{n}\left(m_i - \overline{m}\right)^2\right)^{1/2}} \qquad \begin{array}{l} n, o_i, m_i \text{ defined in (1a)} \\ \overline{o} = \text{mean observed dTWS} \\ \overline{m} = \text{mean modeled dTWS} \end{array} \qquad (5.1b)$$

We use F because it allows us to select parameter sets for which both the timing and amplitude of the modeled seasonal cycle match observations. We define the best parameter set as that which minimizes F; we use it as a surrogate optimum. We perform the exhaustive parameter exploration, which mimics a single-criteria manual calibration.

5.4. Results and Discussion

5.4.1. When given a pseudo-optimum parameter set, which process representation is better?

When given their best parameter sets, SSOIL, DEEP, and AQUIFER are equivalently adept at simulating monthly dTWS in Illinois. For all three models, 22.4 mm ≤ RMSE ≤ 22.7 mm and $r \geq 0.72$ (Figure 5.1). In the ideal case where observations exist and calibration identifies the optimal parameter set, the most computationally efficient model (either SSOIL or AQUIFER) should be used. Single-criteria analysis does not provide sufficient information with which to distinguish the overall performance of the models. Future work will use automatic multi-criteria parameter estimation to further explore the variation in model skill.

5.4.2. When little is known about parameter values, which process representation is best?

In the absence of specific information, modelers often use default parameter sets recommended by model developers. The Taylor diagram [*Taylor,* 2000] in Figure 5.2 shows the performance of both the default and best sets for each of the three models. For all three models, dTWS simulated using the best set has lower variance than observed dTWS, and the improvement over the default set is marginal. The good performance of the default set is not surprising: as one of the few extensive hydrologic datasets in the world, ISWS observations regularly inform LSM development and default parameter estimations.

For most locations, there is a dearth of reliable information about subsurface hydrologic parameters, and we have no way to know whether the default parameter set adequately represents effective parameters. The right panel of Figure 5.2 shows the variations in the consistency of model performance between SSOIL, DEEP, and AQUIFER. Each point represents a single model run in which a unique set of hydraulic parameters was used. Model skill improves as points near the origin. Of the three models, AQUIFER is the least sensitive to choice of parameters. Its robustness likely results from the buffering capacity of the augmented subsurface reservoir. When this buffering mechanism is absent, adequate simulation of dTWS (i.e., realistic regulation of the flux of water into the soil column) depends entirely on accurate assignation of effective parameters. DEEP is slightly less sensitive to faulty parameter values than is SSOIL.

Figure 5.3 compares the empirical cumulative distribution functions of parameter sets of the top-scoring 1% of runs with the distributions of the top-scoring 50%. Because

the distributions of the top 1% differ from those of the top half, the parameters shown in Figure 5.3 are sensitive and merit calibration [*Bastidas et al.,* 1999]. More important, the gentler curvature of AQUIFER's parameter distributions indicates a decrease in sensitivity to percent sand, percent clay, porosity, and maximum rate of subsurface runoff with respect to the other two models, whose cumulative distribution curves are sharper and steeper. Application of a two-sided Kolmogorov–Smirnov test confirms that the difference in sensitivity is statistically significant.

In the foreseeable future, for large model domains, the scientific community is unlikely to be able to confidently assign subsurface hydrologic parameters either by direct observation or by calibration against subsurface hydrologic observations. Decreasing the sensitivity of model output to faulty parameter choices is therefore of utmost practical importance for improving model prediction capability. However, if soil texture properties are the dominant control on regional subsurface hydrologic variation in nature, then AQUIFER's lower sensitivity to parameter values is likely problematic, and a significant increase in data collection and subsequent parameter estimation is warranted.

5.4.3. Does knowledge of parameter ranges guarantee reasonable model output?

The left panel of Figure 5.4 presents the top-scoring 1% of parameter sets for SSOIL and AQUIFER. Within the envelopes created by the top 1%, the best parameter combination is highlighted in black. The right panel presents the scores of all parameter sets for which all values fall within the ranges defined by the envelope created by the top 1%. Note that the right panel does not only show the scores of the "good" runs, which are

clustered close to the origin; it also presents the scores of the runs that used parameter sets that are near those that resulted in the top-scoring 1% of simulations. For instance, "*BAD*" (left panel, dashed line) is a parameter set that, despite of having values within the envelope, performs very poorly (e.g., for SSOIL, RMSE ≈ 0.2 m; for AQUIFER, RMSE ≈ 10 m). Data for DEEP is not shown but is qualitatively similar to that shown for SSOIL. For most parameter sets, AQUIFER performs well. However, we show that there exist parameter sets that are adjacent to top-scoring sets but that result in extremely unrealistic model output. Because of parameter interaction, knowledge of approximate parameter values is insufficient to guarantee realistic simulation of dTWS.

5.5. Conclusions and Implications

When a surrogate optimal parameter set is used, the version of the model with the 3.43-m, 10-layer soil profile; that with the 30-layer, 11.2-m soil profile; and that in which a lumped unconfined aquifer is coupled to the shallow soil profile are equivalently adept at simulating monthly dTWS over the state of Illinois. When knowledge of subsurface hydraulic parameter values is limited, the coupled aquifer model makes CLM significantly less sensitive to errant parameter values; that is, the explicit aquifer representation is the most robust of the three parameterizations. However, knowledge of ranges for individual parameters is insufficient to guarantee realistic simulation of monthly dTWS.

5.6. Acknowledgements

We are grateful to Yasir Kaheil, MinHui Lo, and two anonymous reviewers for their insightful, constructive comments. Thanks to Bob Scott at the ISWS for the soil

moisture data and to Peter Lawrence for the land-surface dataset. Rosero and Gulden thank the NASA GEST Graduate Student Summer Program. Texas Advanced Computing Center supplied computing resources. NASA grants NAG5-10209 and NAG5-12577 and NOAA grant NAO3OAR4310079 provided funding.

Table 5.1. Ranges and distributions of randomly sampled subsurface hydrologic parameters.

Parameter	Range	Distribution
Sand*	5 to 90%	Uniform
Clay*	5 to (100 – [% sand])%	semi-uniform
Porosity	0.01–0.50 m^3m^{-3}	Uniform
e-folding depth of saturated hydraulic conductivity	0.1–100 m	Uniform
Maximum rate of baseflow (rsbmax)	1×10^{-11}–1×10^{-3} m s^{-1}	log uniform
Specific yield†	0.01–0.25	Uniform

* CLM calculates hydraulic conductivity and matric potential as a function of percent sand and percent clay according to the methods of Clapp and Hornberger (1978) and Cosby et al. (1984). Percent silt is 100–(% sand + % clay).

† Specific yield is used only by the AQUIFER runs, not by the SHALLOW or DEEP runs.

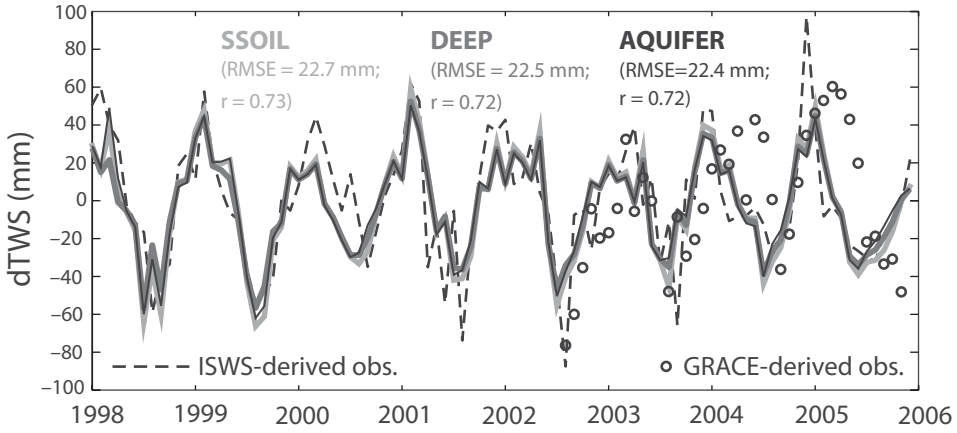

Figure 5.1. Observed monthly dTWS compared with that simulated by each model with its optimal parameter set.

GRACE-derived data (Chen et al., 2006) are shown only for reference; they were not used to score model output.

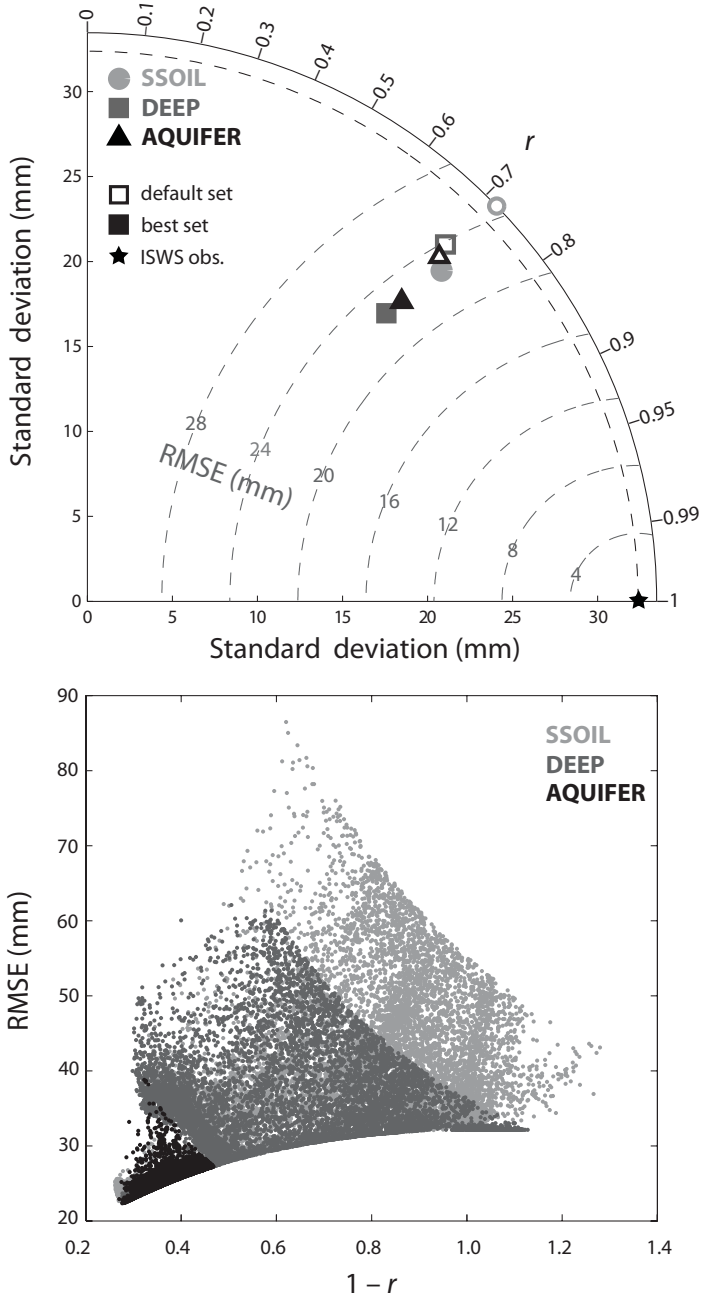

Figure 5.2. Model performance.

A Taylor diagram including scores from runs using default and best parameter sets (left panel) and scores for the top-scoring 50% of runs (right panel). The right panel contains an equal number (11,250) of black, dark gray, and light gray points.

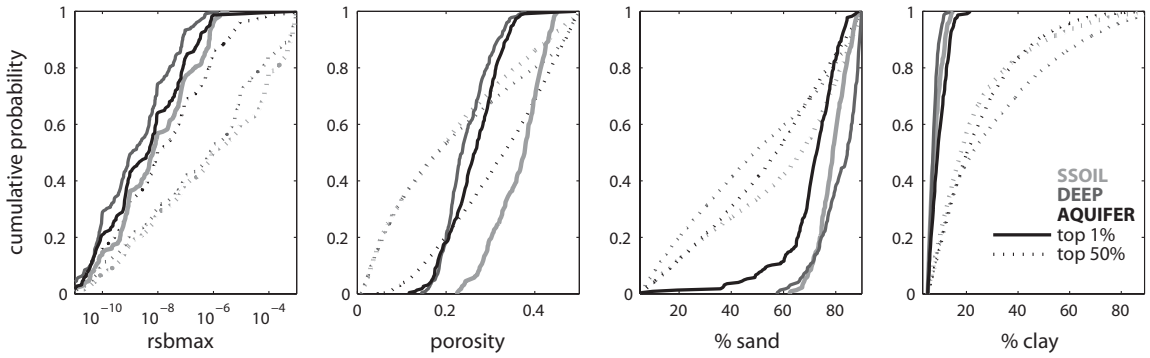

Figure 5.3. Empirical cumulative distribution functions for parameters used in top-scoring 1% and 50% of runs.

Distributions for specific yield and e-folding depth of hydraulic conductivity are not shown.

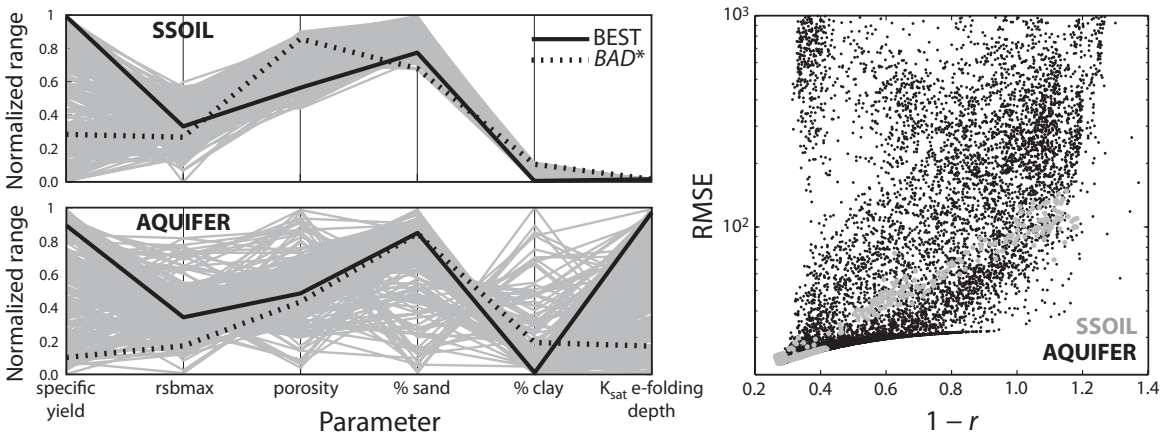

Figure 5.4. Parameter sets corresponding to the top-scoring 1% of runs in normalized parameter space (left panel).

Scores of all the sets that fall within the envelope created by the top 1% (right panel).
*BAD is not a member of the top-scoring 1%; it is the worst-scoring parameter set in the envelope.

Chapter 6: Model performance, robustness, and fitness: *A new way to identify good land-surface models*[5]

6.1. ABSTRACT

Three metrics for rigorous evaluation of land-surface models (LSMs) are introduced. The framework presented here explicitly acknowledges perennial sources of uncertainty in LSM output. The model performance score (ς) quantifies the likelihood that a representative model ensemble will bracket most observations and be highly skilled with low spread. The robustness score (ρ) quantifies the sensitivity of performance to parameter and/or data error. The fitness score (φ) combines performance and robustness, ranking models' suitability for broad application. We demonstrate the use of the metrics by comparing three versions of the Noah LSM. Using time-varying ς for hypothesis testing and model development, we show that representing short-term phenological change improves Noah's simulation of surface energy partitioning and subsurface water dynamics at a semi-humid site. The least complex version of Noah is most fit for broad application. The framework and metrics presented here can significantly improve the confidence that can be placed in LSM predictions.

[5]Substantial portions of this chapter were originally published in Gulden, L. E., E. Rosero, Z.-L. Yang, T. Wagener, and G. Niu (2008b), Model performance, model robustness, and model fitness scores: A new method for identifying good land-surface models, *Geophys. Res. Let.,* 35, L11404, doi:10.1029/2008GL033721. The References section contains full citations for all articles referenced here.

6.2. INTRODUCTION

The increasing reliance of scientists, engineers, and policymakers on the predictions of land-surface models (LSMs) demands more rigorous evaluation of LSM parameterizations. Most LSMs are assessed using limited, localized, often semi-qualitative approaches (e.g., Chen et al., 2007). With few exceptions, model intercomparisons and evaluations of modified parameterizations neglect an assessment of uncertainty that extends beyond simple end-member sensitivity analyses (e.g., Niu et al., 2005). This incomplete approach is due in part to a dearth of observations and in part to evaluation procedures that are no longer state-of-the-art with respect to available computing resources. The development of robust metrics for comprehensive model evaluation is in its infancy (Randall et al., 2007). Here, we present a simple method for increasing the rigor of LSM assessment.

The most straightforward method for assessing an LSM is to evaluate performance at a single site using default parameters (e.g., Henderson-Sellers et al., 1996). LSM performance varies widely when parameters are shifted within reasonable ranges (e.g., Gulden et al., 2007b). At a given site, the parameter set resulting in the best performance may significantly differ from the default. That one model equipped with default parameters does better than another is likely fortuitous. Parameters tend to be effective values, not physical quantities (Wagener and Gupta, 2005). A more thorough evaluation method is to first minimize parameter error by calibrating all models and to then compare model output generated with the best parameter set (e.g., Nijssen and Bastidas, 2005). Using optimal parameters does not represent the way in which LSMs are generally applied; calibration against certain criteria may worsen the simulation of other,

equally important criteria (Leplastrier et al., 2002). After calibration, most equivalently complex models perform equivalently well (e.g., Beven, 2006). Additional methods for LSM evaluation (e.g., the use of neural networks to benchmark LSMs [Abramowitz, 2005]) show promise, but to our knowledge, none has been widely adopted.

Even in the rare case when we can estimate individual parameter ranges, parameter interaction and discontinuous model responses to even small shifts in parameter values limit our confidence in the realism of simulations (e.g., Gulden et al., 2007b; Rosero et al., 2009). The dearth of extensive validation datasets makes this limitation unlikely to soon change. To assess model performance in 'real life' settings, evaluation frameworks such as the one we present here must explicitly acknowledge these sources of uncertainty. Here we treat only parameter uncertainty, but we stress that our framework can and should be applied to incorporate uncertainty in observations.

The dependence of model performance on parameter and forcing error supports the use of a probabilistic approach to evaluate LSMs. To evaluate LSMs, we propose three metrics that harness the information contained in ensemble runs of an individual model. This paper introduces the metrics themselves; Rosero et al. (2009) apply the metrics presented here as part of an in-depth model intercomparison.

6.3. EXAMPLE APPLICATION

To demonstrate the new framework for LSM evaluation, we use an example application. We run three versions of the Noah LSM (Ek et al., 2003) using meteorological forcing data from the 2002 International H_2O Project (IHOP) (LeMone et al., 2007) at sites covered by dry grassland (site 2), semi-arid pasture (site 4) and semi-

humid grassland (site 8) (mean annual precipitation [MAP] = 540, 740, 880 mm y^{-1}, respectively). The standard version of Noah ('STD') is the benchmark against which we evaluate two newer versions: one augmented with a lumped, unconfined aquifer model ('GW') (Niu et al., 2007) and a second augmented with a short-term dynamic phenology module ('DP') that allows leaf area to change in response to environmental variation on daily to seasonal time scales (Dickinson et al., 1998).

Three independent objectives are used to evaluate model performance: 3-hour-running mean evaporative fraction (EF); top 30-cm soil wetness (W_{30}); and 24-hour change in wetness (ΔW_{30}). We define EF as:

$$EF_t = LE_t / \left(LE_t + H_t \right) \tag{6.1}$$

where LE_t and H_t, are, respectively, the latent, and sensible heat flux, averaged over 30-minute time interval t. We compute W_{30} as:

$$W_{30} = \sum_{i=1}^{N_{layers}} \theta_i z_i \Big/ \sum_{i=1}^{N_{layers}} \omega_i z_i \tag{6.2}$$

where θ_i, z_i, and ω_i are, respectively, the volumetric soil moisture, thickness, and porosity of the ith layer of the soil column, which has N_{layers} layers (for the observations, $N_{layers} = 4$; for the models, $N_{layers} = 2$). The 24-hour change in soil-moisture is represented as:

$$\Delta W_{30,t} = W_{30,t} - W_{30,t-47}. \tag{6.3}$$

For each site, we generate two 150-member ensembles: (1) a 'calibrated ensemble,' generated using parameters defined by the Markov Chain Monte Carlo sampler of Vrugt et al. (2003) while simultaneously minimizing five RMSE objectives (LE, H, ground heat flux, 5-cm soil temperature and moisture); and (2) an 'uncalibrated ensemble,' composed

of runs from a subset of 15,000 generated by random sampling of uniform independent parameter distributions; the subset was defined as the group that obtained scores within one standard deviation of the mode for each RMSE (i.e., the most frequent error) (Figure 6.1). In this example, model performance widely varies when parameters are selected within reasonable ranges (Figure 6.1a); after calibration, STD, DP, and GW perform equivalently well (Figure 6.1b). When generating ensembles, for simplicity, we neglected data uncertainty. All realizations used a 30-minute time step to simulate 01/01/2000–06/25/2002. We treated the first 2.5 years of simulation as model spin-up; only the last 45 days of each simulation were scored.

6.4. SCORES OF PERFORMANCE, ROBUSTNESS, AND FITNESS

Before defining a metric that identifies the 'best' model or 'best' parameterization, we first define the characteristics of a good model. Given a representative ensemble, when the LSM is good:

6.1. The ensemble brackets most of the observations.

6.2. The ensemble is centered on the observations.

6.3. The ensemble has low spread when bracketing observations but high spread when not (i.e., the ensemble does not resolutely cling to an incorrect value).

6.4. The model is relatively insensitive to parameters that are not observable and is also not significantly affected by errors in meteorological forcing data.

6.5. The model performance is consistently good (as defined by Descriptors 6.1–6.3) across sites.

Descriptors 6.1–6.3 describe a model that is well suited to a given location; 6.4 and 6.5 describe a model that is robust (Carlson and Doyle, 2002). The ideal model for use over regional and global domains will match all five descriptors.

At time step *t*, the ensemble of the best-scoring model minimizes the performance score, ς_t:

$$\varsigma_t = \left(CDF_{ens,t} - CDF_{obs,t}\right) / \left(1 - \left(CDF_{\overline{obs}+c}\right)\right) \tag{6.4a}$$

where $CDF_{ens,t}$, $CDF_{obs,t}$, and $CDF_{\overline{obs}+c}$ are, respectively, the cumulative distribution functions of the variable simulated by the ensemble of models, of the observation at time *t*, and of all values of $\overline{o}_t$, shifted by arbitrary constant *c* (to prevent division by zero). $\overline{o}_t$ is the mean of all realizations of the observation at time step *t*. When observational uncertainty is not known, for simplicity, ς_t can be expressed in deterministic form as:

$$\varsigma_t = \sum_{i=1}^{N_{ens}} \left| x_{i,t} - o_t \right| \Big/ \sum_{i=1}^{N_{ens}} \left| \overline{o} - c \right| \tag{6.4b}$$

where $x_{i,t}$ is ensemble member *i* at time *t*, o_t is the observation at time *t*, N_{ens} is the number of ensemble members, *c* is an arbitrary constant that is less than all values o_t, and $\overline{o}$ is the mean of the observations. ς_t is lowest (best) when the ensemble brackets most observations, has low spread, and is centered on the observations (Figure 6.2). Figure 6.3 shows the ensemble's time-varying performance and the corresponding ς_t. Table 6.1 demonstrates that ς_t encompasses both the commonly used ensemble spread and skill (e.g., Talagrand et al., 1997).

We can express overall insensitivity to factors that may significantly alter performance (e.g., poorly known parameters or errors in meteorological forcing data) as:

$$\rho = \frac{\left|\bar{\varsigma}_{e1} - \bar{\varsigma}_{e2}\right|}{\bar{\varsigma}_{e1} + \bar{\varsigma}_{e2}} \tag{6.5}$$

where $\bar{\varsigma}_{e1}$ and $\bar{\varsigma}_{e2}$ are the time means of the performance scores for the first and second ensembles, respectively. The two ensembles should significantly differ in the way(s) in which modelers wish to assess robustness. For example, to test robustness with respect to parameter variation, the ensemble members should use parameter sets that come from distinct distributions (as described above); to test robustness with respect to data error, the ensembles should differ in the type or level of noise by which their input data is perturbed.

Given the assumption that spatially varying characteristics of the land surface shape surface-to-atmosphere fluxes and near-surface states (e.g., Dickinson, 1995), we note that there is an inherent contradiction between Descriptors 6.4 and 6.5 above. A tradeoff exists between a model that is completely insensitive to parameter variation and one that consistently does well across sites. A 'compromise ideal' model is insensitive to the parameters that cannot be easily identified; it is at least somewhat sensitive to parameters that are physically realistic (e.g., vegetation type) and when model and measurement scales are similar.

For a given site and criterion, a model's overall fitness score, ϕ, quantifies its suitability for broad application. We define ϕ as:

$$\phi = \bar{\varsigma}\rho \tag{6.6}$$

where $\bar{\varsigma}$ is the mean performance score (Equation 6.4) of the most representative ensemble and ρ is robustness (Equation 6.5). The more sites and criteria used, the more confident we can be of model performance (Descriptor 6.5). The best model (*m*) from a

set of M models tested across N sites minimizes $\frac{1}{N}\sum_{i=1}^{N}\phi_{i,m}$. For fair cross-criterion comparison, modelers should first rank ϕ for a given criterion and should then compare average fitness rankings of the models.

Figures 6.4a and 6.4b show how this framework can be used to improve the physical structure of LSMs. DP, which is different from benchmark model STD only in that it allows leaf area index to vary over short time scales in response to environmental variation, consistently better simulates EF at semi-humid site 8 (Figure 6.4a), a result that is consistent with the hypothesis that short-term phenology can alter surface energy partitioning (see Rosero et al. [2009] for detailed analysis). DP tends to have the best ς_t when simulating 24-hour change in wetness (Figure 6.4b, Table 6.2). When used with other model output characteristics (e.g., anomalies, bias), ς_t helps determine when and where the model is most likely to succeed or fail. It can also be used across criteria (Figure 6.4c); the combination of ς and ρ yields an assessment of overall model fitness (Figure 6.4d and Table 6.2). Overall, GW performs best but is less robust than STD. STD is the fittest model, but tradeoffs in fitness between criteria make all models equivalently fit at the most-humid site 8 (Table 6.2).

6.5. SUMMARY AND IMPLICATIONS

We introduce three metrics for rigorous and realistic evaluation of LSM performance within a framework that explicitly acknowledges perennial sources of LSM output uncertainty (e.g., sensitivity of output to parameters that are impossible to specify and to errors in meteorological forcing data). The model performance score (ς) quantifies

the likelihood that a representative model ensemble will be highly skilled with low spread; the robustness score (ρ) quantifies the sensitivity of model performance to changes in parameters (as shown in the example here) and/or perturbations to meteorological forcing. Our framework treats the relative insensitivity of an LSM to both parameter variability and to forcing error as beneficial characteristics in the face of the less-than-perfect settings in which LSMs are applied. The fitness score (ϕ) combines the concepts of good performance and robustness and is used to rank models' suitability for broad application.

The use of the metrics is demonstrated using three versions of the Noah LSM to simulate summer in Oklahoma. Our example shows that the least complex version of Noah is most fit for broad application. We use the time-varying ς (a tool for model evaluation and development) to show that allowing leaf area index to vary on short time scales improves Noah's simulation of surface energy partitioning and subsurface water dynamics at the semi-humid site. Standard computational resources are now such that the presented framework can be applied to several models (or a single model with candidate parameterizations) and numerous flux tower sites as a means for more thorough and informative model evaluation: on a single 2.66 GHz processor, to run one model for 2.5 years 15,000 times (as we did for each Monte Carlo sampling) required less than 2 hours of computing time.

Researchers are often quick to assume that a model is performing well because it is more complex (i.e., has more parameters), a characteristic that is often equated with increased physical realism. This method for evaluation should not be used to 'prove' that one representation is more physically realistic than another. Although it is likely that

improved conceptual realism will improve model performance, the converse is not necessarily valid. Regardless, the models examined here have so many degrees of freedom that it is difficult to parse whether strong model performance is the result of cancelling errors or the result of physical correctness; however, the performance and fitness scores presented here allow for hypothesis testing and model development that gives a realistic treatment to uncertainty. Because the results are obtained using ensemble simulations, we can be more confident that this improvement is indeed the result of the altered model structure and is not the simple result of a lucky guess of parameters.

Land-surface modelers are unlikely to ever know the 'right' parameters for any site on which their models are applied; they will not be able to eliminate error in meteorological forcing data. Evaluation of the performance of LSMs must take a realistic view of these limitations. The metrics proposed here enable more objective comparison of LSM performance in a framework that more accurately represents the ways in which LSMs are applied. Use of this framework and metrics will strengthen modelers' conclusions and will improve confidence in LSM predictions.

6.6. ACKNOWLEDGEMENTS

We thank the Texas Advanced Computing Center, NOAA (grant NA07OAR4310216), the NSF, the NWS/OHD, and the Jackson School of Geosciences; we are also grateful for the insightful comments of two anonymous reviewers.

Table 6.1. Ensemble spread, skill, and performance score for EF simulation at IHOP Site 2 at 2:00 PM local time on Julian day 161 of 2002.*

	Spread†		Skill§		Performance score (ς)	
	Calibrated	Uncalibrated	Calibrated	Uncalibrated	Calibrated	Uncalibrated
STD	0.00207	0.0037	0.00226	0.0047	0.114	0.154
DP	0.0105	0.0113	0.0109	6.07e−7	0.232	0.183
GW	0.00291	0.00504	9.22e−6	0.00293	0.0815	0.152

* See also Figure 2.

†Ensemble spread (π_t) is $\pi_t = \frac{1}{N_{ens} - 1} \sum_{i=1}^{N_{ens}} \left(x_{i,t} - \bar{x}_t \right)^2$, where $\bar{x}_t$ is the ensemble mean at time *t,* $x_{i,t}$ is the *i*th ensemble member at time *t*, and N_{ens} is the number of ensembles.

§Ensemble skill (κ_t) is $\kappa_t = \left(\bar{x}_t - o_t \right)^2$, where o_t is the observed value at time *t.*

Table 6.2. Model fitness and average ranking.*

	Criterion	STD	DP	GW
Site 2	EF	0.0085 (1)	0.0278 (3)	0.0214 (2)
	ΔW_{30}	0.0329 (2)	0.0635 (3)	0.0155 (1)
	W_{30}	0.0030 (1)	0.1485 (3)	0.0424 (2)
	Site-mean rank	1.33	3	1.67
Site 4	EF	0.0041 (1)	0.0252 (3)	0.0164 (2)
	ΔW_{30}	0.1997 (2)	0.0901 (1)	0.6172 (3)
	W_{30}	0.0004 (1)	0.0085 (2)	0.0256 (3)
	Site-mean rank	1.33	2	2.67
Site 8	EF	0.0346 (2)	0.0241 (1)	0.0546 (3)
	ΔW_{30}	0.5246 (3)	0.3220 (2)	0.2905 (1)
	W_{30}	0.0235 (1)	0.0497 (3)	0.0403 (2)
	Site-mean rank	2	2	2
	Mean rank	1.55	2.33	2.11
	Variance of rank	0.53	0.75	0.61

* The rank of model fitness score among the cohort for each given site and criterion is shown in parentheses.

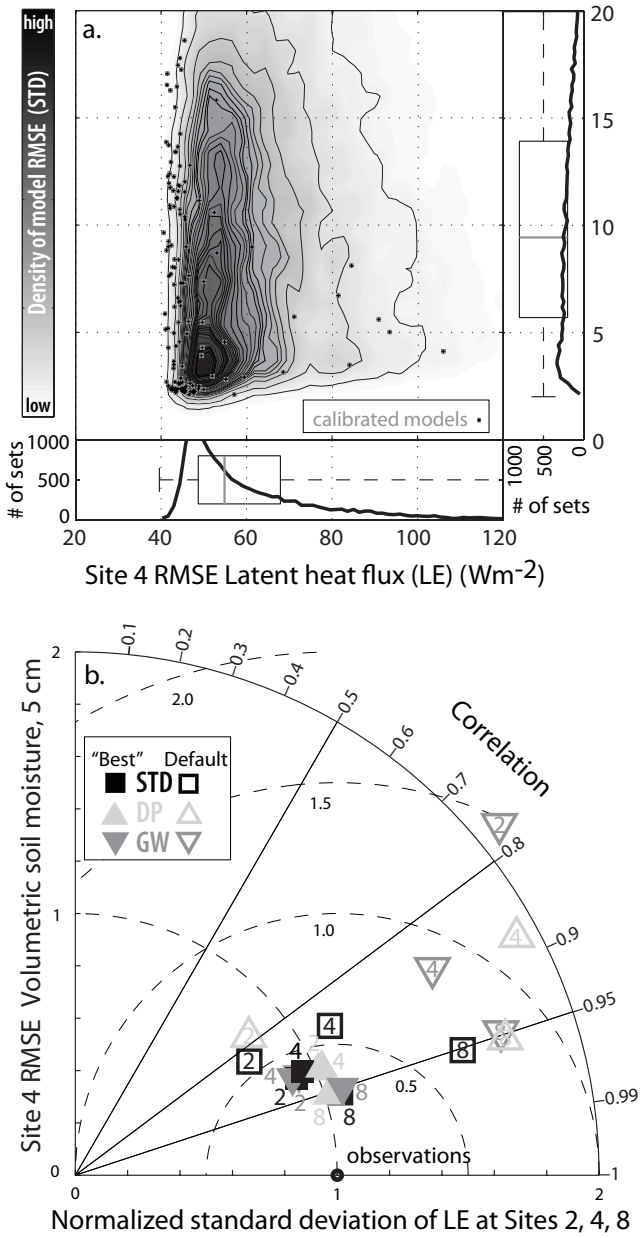

Figure 6.1. The range of scores obtained by varying effective parameters.

(a) RMSE for 5-cm volumetric soil moisture and LE for 15,000 runs of STD at IHOP site 4. Each simulation used unique parameter sets randomly selected from uniform distributions. Also shown are calibrated-model scores. (b) Taylor diagram of LE simulated by DP, STD, and GW at 3 IHOP sites (2, 4, 8).

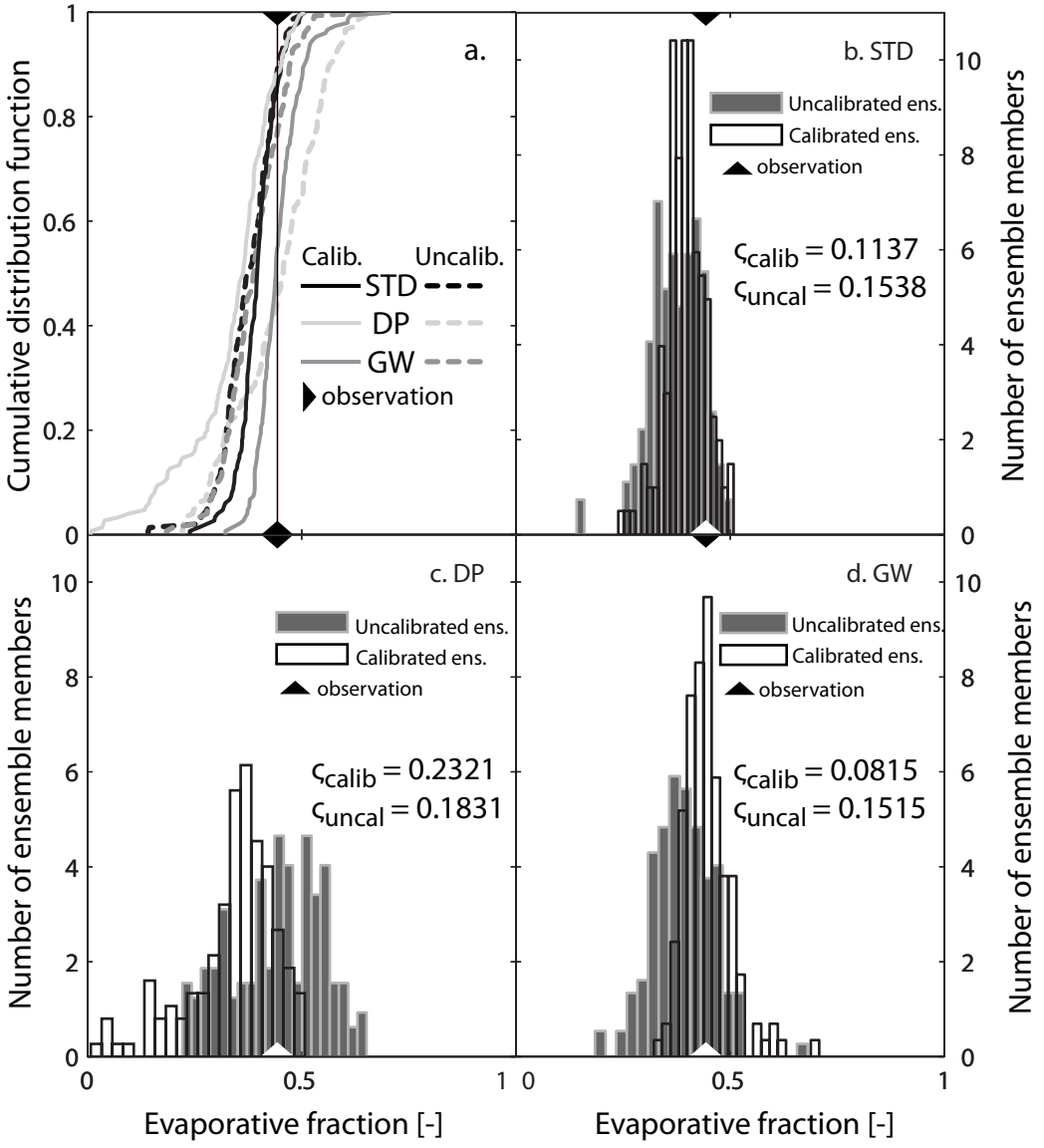

Figure 6.2. Distribution of EF simulated for Site 2 at 2:00 PM on Julian day 161 (2002).

(a) CDFs of the calibrated and uncalibrated ensembles; (b)–(d) Performance scores. The performance score improves as the model is better described by Descriptors 6.1–6.3 (see text).

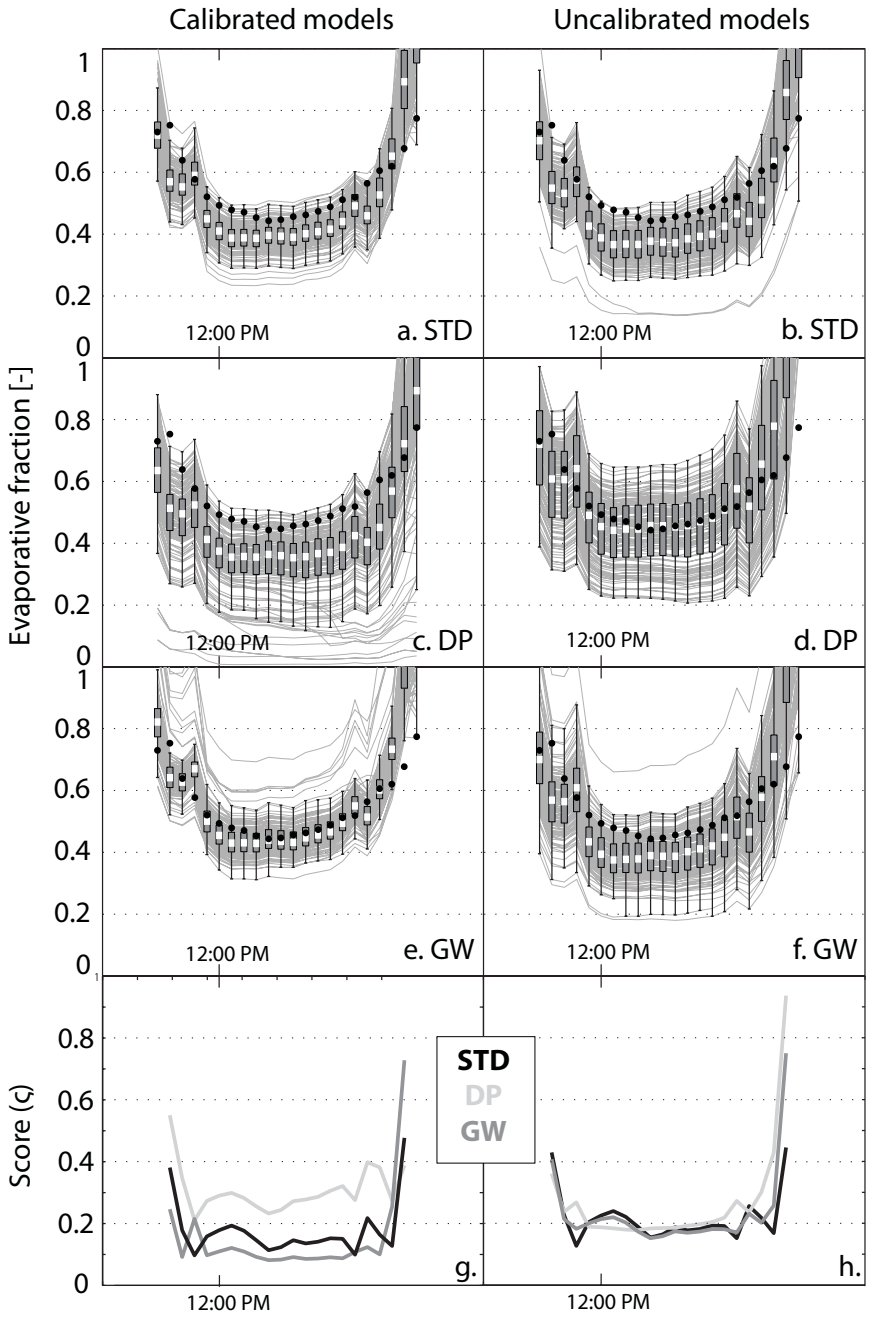

Figure 6.3. Relation between time-varying ensemble simulations of EF and time-varying performance score (ς).

EF simulated by each ensemble is shown for Site 2 for Julian day 161 (2002). Black circles are observations; gray lines are individual ensemble members; white bars are ensemble mean; gray bars are the ensemble interquartile range.

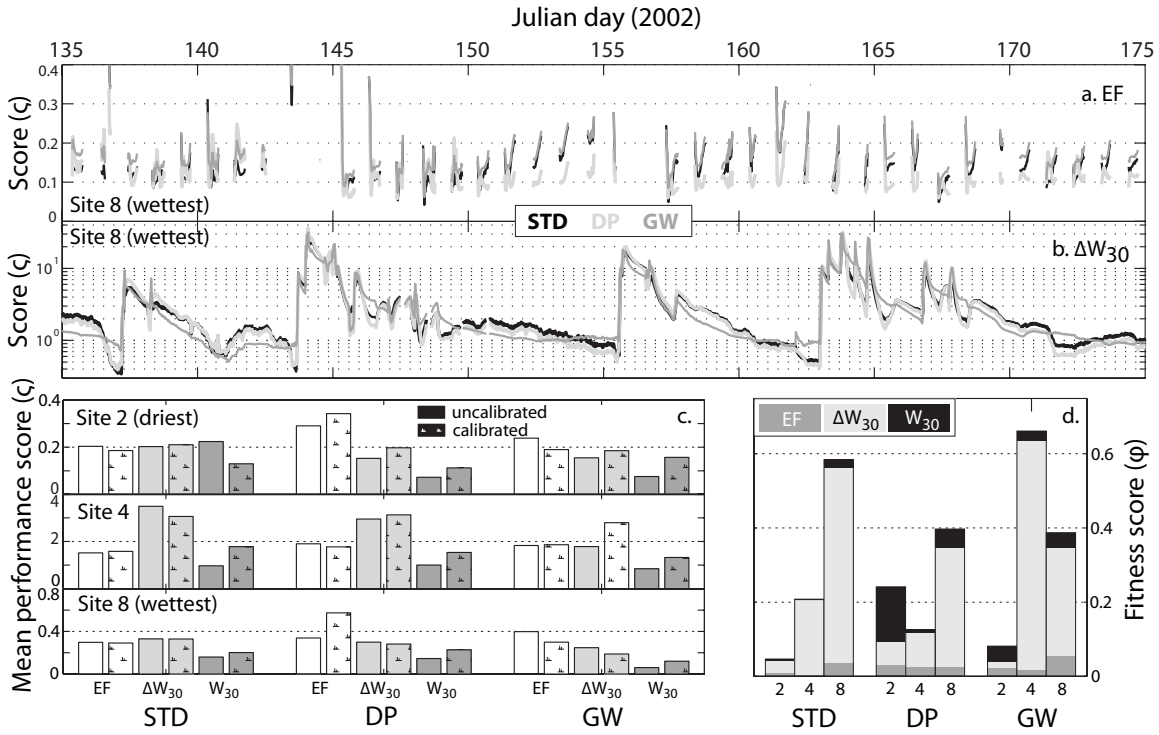

Figure 6.4. Utility of metrics.

Time-varying performance score (ς) of STD, DP, and GW for (a) EF and (b) 24-hour change in wetness (ΔW30). (c) Model performance when simulating EF, ΔW30, and ΔW30; change between the calibrated and uncalibrated ensemble performance scores indicates model robustness. (d) Model fitness scores for each model, criterion, and site.

Chapter 7: Conclusions

7.1. SUMMARY AND IMPLICATIONS

The research presented in Chapters 2, 3, and 4 addressed science questions related to the emission of biogenic volatile organic compounds while exploring different contributors to uncertainty in LSM simulations (i.e., input data uncertainty and process uncertainty). Research presented in Chapter 5 addressed the interplay between a model's conceptual realism and the model's sensitivity to parameter uncertainty. Chapter 6 presented a framework and associated metrics that allow researchers to quantify, in time, typical model performance given levels of parameter uncertainty or input data uncertainty that are characteristic of the settings in which the models are normally applied. Most of the conclusions drawn from the research completed for this dissertation can be extrapolated to similar land-surface processes or similar challenges in environmental modeling. All research presented here in some way informs the confidence that can be placed in the simulations of LSMs.

Chapter 2 described a process to develop a species-based land-cover dataset and presented a bottom-up classification method for deriving species-based biogenic emissions factors for the land-cover categories. The inherent biogenic emission capacity of the Texas landscape that was computed using the simplified dataset and species-based emissions factors was well correlated (r = 0.89) with that computed using the original, complex dataset containing 600+ plant species and corresponding emissions factors. The results showed that when LSMs are equipped with species-based emission factors, they

have the capacity to realistically represent the spatial distribution of biogenic emissions. The method presented can be applied to other regions and may be extrapolated to condense similar information (e.g., CO_2 uptake rates) for use in LSMs or other environmental models. The research provides evidence that LSMs, although simplified versions of nature that are subject to uncertainty, have the capacity to represent complex and spatially heterogeneous processes with first-order accuracy.

Chapter 3 showed that biogenic emissions predictions made by LSMs are particularly sensitive to the vegetation characteristics in the input land-cover dataset, which, like many types of data used to drive LSMs, are not well constrained by observations. When simulating biogenic emissions in Texas, varying only the vegetation distribution used to initialize the LSM causes emissions estimates to vary, on average, by a factor of 3. Biogenic emissions estimates are most sensitive to variation between the spatial distributions of plant types that are specified by land-cover datasets. Variation in simulated biogenic emissions that stems from uncertainty in land-cover is highest in eastern and central Texas, where biogenic emissions are highest and where tropospheric ozone pollution is a significant concern. The results presented in Chapter 3 raise a red flag regarding the confidence that can be placed in LSM predictions of biogenic emissions, which are shown to be highly dependent on uncertain land-cover data. This caveat is especially relevant when one reviews projections for the mid-to-distant future, when land-cover composition may not match the current-day landscape. The uncertainty inherited from land-cover data transfers to the projections made by models that are coupled to LSMs (e.g., atmospheric chemistry models within numerical weather prediction models).

Chapter 4 addressed the extent to which interannual variation in biomass, a process that had previously been omitted from LSMs, alters biogenic emissions. Our results are consistent with the hypothesis that interannual variation in LAI is as important as or more important than interannual variation in shortwave radiation and leaf temperature to the control of biogenic emissions. More-humid regions with less interannual variation in precipitation show lower year-to-year variation in biogenic emissions: as a grid cell's mean biogenic emission flux increases, the mean-normalized standard deviation of that flux tends to decrease. Especially when interannual variation in biogenic emissions is important to researchers or policymakers (e.g., when predicting future changes in the number of days on which urban air exceeds acceptable ozone concentrations), year-to-year change in biomass should be represented within models. In the broader context of environmental modeling, the research presented in Chapter 4 serves as a reminder for the need to continually evaluate the suitability of model process representations to the task for which the model is being used.

LSMs are often developed and evaluated at idealized, data-rich sites but are then applied in regions where data are either sparse or unavailable. The research presented in Chapter 5 evaluated three different ways of parameterizing subsurface hydrology in LSMs and explored how well each parameterization performed under increasing parameter uncertainty. When given their optimal parameter sets, all three versions of the LSM ((1) with a standard shallow, 10-layer soil profile; (2) with a deeper, many layered soil column; (3) with an explicit lumped aquifer model) perform equivalently well in simulating monthly change in terrestrial water storage (i.e., there is equifinality among models). The most conceptually realistic model (in this case, the model with an explicit

aquifer) is significantly less sensitive to errant parameter values: that is, when we know nothing about 'correct' parameter values, the most conceptually realistic model is most likely to perform reasonably well. However, even when using the most conceptually realistic model, because of parameter interaction, knowledge of ranges for individual parameters is insufficient to guarantee realistic simulation of monthly change in terrestrial water storage. All three major findings of Chapter 5 have broad implications for model development. Equifinality renders the identification of 'better' models and 'good' parameterizations a more labor-intensive task than simply comparing single runs of each model, a practice that is the current standard in model development (e.g., Gulden et al., 2007a). Equifinality, combined with the recognition that LSM simulations are sensitive to the values of many uncertain parameters, brings into question the strength of the conclusions derived from model intercomparisons (e.g., Henderson-Sellers et al., 1996). That the most conceptually realistic model tends to be most reliable despite parameter uncertainty increases our confidence in the reliability of process-based representations; however, the remaining capacity for significant model error is a caveat against overconfidence in the performance of models believed to be conceptually accurate.

In an effort to make more quantitative the analysis of LSMs in ways that are consistent with the conditions in which they are actually used, in Chapter 6, I presented a framework and associated metrics for evaluation of LSMs. The framework explicitly acknowledges perennial sources of uncertainty in LSM predictions (e.g., parameter uncertainty, forcing-data uncertainty). The model performance score quantifies the likelihood that an ensemble that is representative of model performance will bracket most

observations, will be highly skilled, and will have low spread. The robustness score quantifies the sensitivity of model performance to errors in parameter choice or input/forcing data. The fitness score combines performance and robustness, ranking models' suitability for broad application. An important feature of the proposed framework is its capacity to show how model performance changes in time, which allows for the metrics to be used as diagnostic tools in model development (Rosero et al., 2009). The framework and metrics presented here have the potential to significantly improve the quantification of the confidence that can be placed in LSM simulations. Quantitatively and computationally intensive testing and evaluation of existing models will likely become standard practice as policymakers and other scientists demand more rigorous assessment of model predictive and explanatory capabilities. If the proposed or a similarly motivated framework (e.g., Abramowitz et al., 2008) is employed as standard practice in model evaluation and development, scientists who use LSM results will be able to provide policymakers with more reliable estimates of predictive uncertainty, which will likely contribute to increased trust and goodwill between the policymaking and scientific communities. Modelers using the framework are able to draw conclusions about model parameterizations and model function that do not depend on the use of error-free input data, error-free evaluation data, or ideal model parameters (all three of which are often difficult, if not impossible, to obtain).

7.2. FUTURE WORK

Questions remain for all segments of my dissertation research. With regard to biogenic emissions, the next generation of LSMs that are coupled to models of the

atmosphere will include representations of the secondary organic aerosols that are reaction products of biogenic volatile organic compounds. Secondary organic aerosols mediate the effect of biogenic emissions on could formation, Earth's radiative balance, and the global carbon cycle. Using data such as that presented here, by employing the next generation of models, scientists will be able to directly address science questions regarding land-surface−atmosphere feedbacks that are mediated by biogenic emissions and their reaction products. For example: Do changes in biogenic emissions mediated by land-cover change drive changes in precipitation patterns? Does climate change alter the sensitivity of cloud formation to biogenic emissions? Do biogenic emissions patch holes in the calculated global carbon budget? In all explorations of the role of biogenic emissions in shaping Earth system processes, an increase in the availability of ecosystem-level validation data will be a welcome asset for increasing our ability to quantify confidence in model-predicted biogenic emissions estimates.

Further exploration of the relationship between model conceptual realism and model reliability is warranted both in the case of the LSM tested in Chapter 5 but also for other LSMs and LSM parameterizations. For example: Does the conclusion that the most conceptually realistic representation of subsurface hydrology is least sensitive to errant parameter choices hold for biomes different than the mid-latitude continental climate of Illinois? Do similar conclusions hold for other model parameterizations (e.g., evaporation from bare soil, sensible heat flux, etc.)? How can we account for errors in one part of an LSM while testing the conceptual realism of a different component?

In continuing to develop the framework presented in Chapter 6, a vital next step is the exploration of the sensitivity of the metrics to the number of runs used to compute

them. It is also worthwhile to investigate the relationship between model fitness and the transferability of parameters between sites and to quantify how model sensitivity that stems from parameter interactions affects model robustness. Researchers need to explore under what conditions a model's sensitivity to parameters (or input data) is beneficial. The framework can be immediately applied in several ways. The metrics presented in Chapter 6 give model developers a tool with which to explore how model performance varies in time and can be applied to any hypothesis testing that employs models. The metrics can serve as good tools with which to explore the relationship between model complexity, model performance, and model physical realism. The framework or a similarly robust model evaluation technique should be employed as standard practice in model sensitivity analysis: i.e., which source of uncertainty affects a given model (or module's) performance the most? Given sufficient data for evaluation, the framework presented in Chapter 6 can be readily applied to distributed models, providing a spatially varying assessment of model performance.

References

Abbot, D. S., et al. (2003), Seasonal and interannual variability of North American isoprene emissions as determined by formaldehyde column measurements from space, *Geophys. Res. Lett.*, 30 (17), 1886, doi:10.1029/2003GL017336.

Abramowitz, G (2005), Towards a benchmark for land surface models, *Geophys. Res. Lett.*, doi:10.1029/2005GL024419.

Abramowitz G., R. Leuning, M. Clark, A. Pitman A (2008), Evaluating the performance of land-surface models, *J. Climate*, 21, 5468-5481.

Altshuller, P. (1991), The production of carbon monoxide by the homogeneous NOx-induced photooxidation of volatile organic compounds in the troposphere, *J. Atm. Chem.*, 13, 155-182.

Andreae, M. O. and P. J. Crutzen (1997), Atmospheric aerosols: Biogeochemical sources and role in atmospheric chemistry, *Science,* 276, 1052-1058.

Andreaevc, M. O., and P. J. Crutzen (1997), Atmospheric aerosols: biogeochemical sources and role in atmospheric chemistry, *Science,* 276, 1052-1058.

Ashmore, M. R. (2005), Assessing the future global impacts of ozone on vegetation, *Plant Cell Env.,* 28, 949-964.

Bastidas, L. A., et al. (1999), Sensitivity analysis of a land surface scheme using multicriteria methods, *J. Geophys. Res.,* 104, 19481-19490.

Beck, M. B. (1987), Water quality modeling: a review of the analysis of uncertainty, *Water Resour. Res.,* 23(8), 1393-1442.

Beven K. (2006), A manifesto for the equifinality thesis, *J. Hydrol,* 320: 18-36.

Bonan, G. B., and S. Levis (2006), Evaluating aspects of the community land and atmosphere models (CLM3 and CAM3) using a Dynamic Global Vegetation Model, *J. Clim.,* 19, 2290-2301.

Bonan, G. B., et al. (2002a), The land surface climatology of the community land model coupled to the NCAR community climate model. *J. Clim.,* 15, 3123-3149.

Bonan, G. B., S. Levis, L. Kergoat, and K. W. Oleson (2002b), Landscapes as patches of plant functional types: An integrating concept for climate and ecosystem models. *Global Biogeochem. Cycles,* 16, doi:10.1029/2000GB001360.

Byun, D. W., et al. (2005), Estimation of biogenic emissions with satellite-derived land use and land cover data for air quality modeling of Houston-Galveston ozone nonattainment area. *J. Env. Mgmt.,* 75, 285-301.

Carlson J.M. and J. Doyle (2002), Complexity and robustness, P Natl Acad Sci USA 99: 2538-2545 Suppl. 1 FEB 19 2002.

Chameides, W. L., et al. (1988), The role of biogenic hydrocarbons in urban photochemical smog - Atlanta as a case-study, *Science,* 241, 1473-1475.

Changnon, S, et al. (1988), Relations between precipitation and shallow groundwater in Illinois, *J. Clim.,* 1, 1239-1250.

Chen F., et al. (2007), Description and Evaluation of the Characteristics of the NCAR High-Resolution Land Data Assimilation System, *J. Appl. Met. Climatol*., 46, 694-713, DOI: 10.1175/JAM2463.1

Chen, F., and J. Dudhia (2001), Coupling an advanced land surface-hydrology model with the Penn State-NCAR MM5 modeling system. Part I: Model implementation and sensitivity. *Mon. Weath. Rev.,* 129, 569-585.

Chen, J. L., C. R. Wilson, and K. W. Seo (2006), Optimized smoothing of gravity recovery and climate experiment (GRACE) time-variable gravity observations, *J. Geophys. Res.,* 111(B6): Art. No. B06408.

Claeys, M., B. Graham, G. Vas, W. Wang, R. Vermeylen, V. Pashynska, J. Cafmeyer, P. Guyon, M. O. Andreae, P. Artaxo, and W. Maenhaut (2004), Formation of secondary organic aerosols through photooxidation of isoprene. *Science,* 303, 1173-1176.

Claeys, M., et al. (2004), Formation of secondary organic aerosols through photooxidation of isoprene. *Science,* 303, 1173-1176.

Clapp, R. B., and G. M. Hornberger (1978), Empirical equations for some soil hydraulic-properties, *Wat. Resour. Res.,* 14, 601-604.

Cleveland, C. C., and J. B. Yavitt (1998), Microbial consumption of atmospheric isoprene in a temperate forest soil, *Appl. Environ. Microbiol.,* 64, 172-177.

Collatz, G. J., C. Grivet, J. T. Ball, and J. A. Berry (1991), Physiological and environmental-regulation of stomatal conductance, photosynthesis and transpiration—A model that includes a laminar boundary-layer, *Agric. For. Meteorol.,* 54, 107-136.

Collatz, G. J., M. Ribas-Carbo, and J. A. Berry (1992), Coupled photosynthesis-stomatal conductance model for leaves of C4 plants, *Aust. J. Plant Physiol.,* 19, 519-538.

Cosby, B. J., et al. (1984), A statistical exploration of the relationships of soil-moisture characteristics to the physical-properties of soils, *Wat. Resour. Res.,* 20, 682-690.

Cosgrove, B. A., et al. (2003), Real-time and retrospective forcing in the North American Land Data Assimilation System (NLDAS) project, *J. Geophys. Res.,* 108(D22), 8842, doi:10.1029/2002JD003118.

Dickinson, R. E. (1983), Land surface processes and climate—Surface albedos and energy balance, *Adv. Geophys.,* 25, 305-353.

Dickinson, R. E., M. Shaikh, R. Bryant, and L. Graumlich (1998), Interactive canopies for a climate model, *J. Clim.,* 11, 2823-2836.

Dickinson, R.E. (1995), Land atmosphere interaction. U.S. National Report to International Union of Geodesy and Geophysics 1991-1994, *Rev. Geophys.,* Supplement, 917-922.

Dougherty, R.L., Bradford, J.A., Coyne, P.I., and Sims, P.L. 1994. Applying an empirical model of stomatal conductance to three C-4 grasses. *Agric. For. Meteor.* 67:269-290.

Ek, M. B., K. E. Mitchell, Y. Lin, E. Rogers, P. Grunmann, V. Koren, G. Gayno, and J. D. Tarpley (2003), Implementation of Noah land surface model advances in the National Centers for Environmental Prediction operational mesoscale Eta model, *J. Geophys. Res.,* 108(D22), 8851, doi:10.1029/2002JD003296.

Fan Y., G. Miguez-Macho, C. P. Weaver, R. Walko, A. Robock (2007), Incorporating water table dynamics in climate modeling: 1. Water table observations and equilibrium water table simulations, J. Geophys. Res., 112, D10125, doi:10.1029/2006JD008111.

Farquhar, G.D., von Caemmerer, S., and Berry, J.A. (1980), A biochemical model of photosynthetic CO2 assimilation in leaves of C species, *Planta,* 149, 78-90.

Fuentes, J. D., and D. Wang (1999), On the seasonality of isoprene emissions from a mixed temperate forest, *Ecol. Appl.,* 9, 1118-1131.

Fuentes, J. D., et al. (1995), Modeled and field-measurements of biogenic hydrocarbon emissions from a canadian deciduous forest, *Atm. Env.,* 29, 3003-3017.

Funk, J. L., C. G. Jones, D. W. Gray, H. L. Throop, L. A. Hyatt, and M. T. Lerdau (2005), Variation in isoprene emission from Quercus rubra: Sources, causes, and consequences for estimating fluxes. *J. Geophys. Res.,* 110, D04301, doi:10.1029/2004JD005229.

Funk, J. L., C. G. Jones, C. J. Baker, H. M. Fuller, C. P. Giardina, and M. T. Lerdua (2003), Diurnal variation in the basal emission rate of isoprene, *Ecol. Appl.,* 13, 269-278.

Geron, C., et al. (2001), Isoprene emission capacity for US tree species, *Atm. Env.,* 35, 3341-3352.

Geron, C., et al. (2000), Temporal variability in the basal isoprene emission factor, *Tree Physiol.,* 20, 799-805.

Gholz, H. L. (1982), Environmental limits on above-ground net primary production, leaf-area, and biomass in vegetation zones of the Pacific northwest, *Ecology,* 63, 469-481.

Grell, G. A., J. Dudhia, and D. R. Stauffer (1995), A description of the fifth-generation Penn State/NCAR Mesoscale Model. NCAR Technical Note, NCAR-TN 398+STR. Accessed October 19, 2005 at http://www.mmm.ucar.edu/mm5/ documents/mm5-desc-pdf/cover.pdf.

Grell, G. A., S. E. Peckham, R. Schmitz, S. A. McKeen, G. Frost, W. C. Skamarock, and B. Eder (2005), Fully coupled "online" chemistry within the WRF model, *Atm. Env.,* 39, 6957-6975.

Grier, C. C., and S. W. Running (1977), Leaf area of mature northwestern coniferous forests - relation to site water-balance, *Ecology,* 58, 893-899.

Guenther, A. (1997), Seasonal and spatial variations in natural volatile organic compound emissions, *Ecol. Appl.,* 7, 34-45.

Guenther, A. (2002), The contribution of reactive carbon emissions from vegetation to the carbon balance of terrestrial ecosystems, *Chemosphere,* 49, 837-844.

Guenther, A. B., R. K. Monson, and R. Fall (1991), Isoprene and monoterpene emission rate variability: Observations with eucalyptus and emission rate algorithm development, *J. Geophys. Res.,* 96(D6), 10799-10808, 10.1029/91JD00960.

Guenther, A., C. N. Hewitt, D. Erickson, R. Fall, C. Geron, T. Graedel, P. Harley, L. Klinger, M. Lerdau, W. A. McKay, T. Pierce, B. Scholes, R. Steinbrecher, R. Tallamraju, J. Taylor, and P. Zimmerman (1995), A global model of natural volatile organic-compound emissions. *J. Geophys. Res.,* 100, 8873-8892.

Guenther, A., et al. (2006), Estimates of global terrestrial isoprene emissions using MEGAN (Model of Emissions of Gases and Aerosols from Nature), Atmospheric Chemistry and Physics, 6, 3181-3210.

Guenther, A., P. Zimmerman, and M. Wildermuth (1994), Natural volatile organic-compound emission rate estimates for United-States woodland landscapes. *Atm. Env.,* 28, 1197-1210.

Gulden, L. E. and Z. L. Yang (2006), Development of species-based, regional emission capacities for simulation of biogenic volatile organic compound emissions in land-surface models: An example from Texas, USA, *Atm. Env.,* 40, 1464-1479.

Gulden, L. E., Z.-L. Yang, and G.-Y. Niu (2007a), Interannual variation in biogenic emissions on a regional scale, *J. Geophys Res.,* 112, D14103, doi:10.1029/2006JD008231.

Gulden, L.E., E. Rosero, Z. Yang, M. Rodell, C. S. Jackson, G. Niu, P. J.-F. Yeh, and J. Famiglietti (2007b), Improving land-surface model hydrology: Is an explicit aquifer model better than a deeper soil profile? *Geophys. Res. Lett.,* 34, L09402, doi:10.1029/2007GL029804.

Gulden, L. E., Z.-L., Yang, and G.-Y. Niu, (2008a), Sensitivity of biogenic emissions simulated by a land-surface model to land-cover representations, *Atm. Env.,* doi:10.1016/j.atmosenv.2008.01.045.

Gulden, L. E., E. Rosero, Z.-L. Yang, T. Wagener, and G. Niu (2008b), Model performance, model robustness, and model fitness scores: A new method for identifying good land-surface models, *Geophys. Res. Lett.,* 35, L11404, doi:10.1029/2008GL033721.

Hakola, H., et al (2003), Seasonal variation of VOC concentrations above a boreal coniferous forest, *Atm. Env.,* 37, 1623-1634.

Hanna, S. R., A. G. Russell, J. G. Wilkinson, J. Vukovich, and D. A. Hansen (2005), Monte Carlo estimation of uncertainties in BEIS3 emission outputs and their effects on uncertainties in chemical transport model predictions, *J. Geophys. Res.,* 110, D01302, doi:10.1029/2004JD004986.

Hansen, M. C., et al. (2000), Global land cover classification at 1km spatial resolution using a classification tree approach, *Internat. J. Rem. Sens.,* 21, 1331-1364.

Henderson-Sellers A., K. McGuffie, and A. J. Pitman (1996), The project for intercomparison of land-surface parameterization schemes (PILPS): 1992 to 1995, *Clim. Dyn.,* 12, 849-859.

Henderson-Sellers, A. (1993), A factorial assessment of the sensitivity of the BATS land-surface parameterization scheme, *J. Clim.,* 6, 227-247.

Hogue T.S., L. A. Bastidas, H. V. Gupta, and S. Sorooshian (2006), Evaluating model performance and parameter behavior for varying levels of land surface model complexity, *Water Resour. Res.,* 42, W08430, doi:10.1029/2005WR004440

Hollinger, S. E., and S. A. Isard (1994), A soil moisture climatology of Illinois, *J. Clim.,* 7, 822-833.

Junquera, V., M. M. Russell, W. Vizuete, Y. Kimura, and D. Allen (2005), Wildfires in eastern Texas in August and September 2000: Emissions, aircraft measurements, and impact on photochemistry, *Atm. Env.,* 39, 4983-4996.

Kanakidou, M., et al. (2000), Human-activity-enhanced formation of organic aerosols by biogenic hydrocarbon oxidation, *J. Geophys. Res.,* 105(D7), 9243-9254, 10.1029/1999JD901148.

Karlik, J. F., et al. (2002), A survey of California plant species with a portable VOC analyzer for biogenic emission inventory development, *Atm. Env.,* 36, 5221-5233.

Kavouras, I. G., N. Mihalopoulos, and E. G. Stephanou (1998), Formation of atmospheric particles from organic acids produced by forests, *Nature,* 395, 683-686.

Kim Y., G. Wang (2005), Modeling seasonal vegetation variation and its validation against Moderate Resolution Imaging Spectroradiometer (MODIS) observations over North America, *J. Geophys. Res.,* 110, D04106, doi:10.1029/2004JD005436.

Kinnee, E., C. Geron, and T. Pierce (1997), United States land use inventory for estimating biogenic ozone precursor emissions, *Ecol. Appl.,* 7(1), 46-58.

Koster, R. D., M. J. Suarez, A. Ducharne, M. Stieglitz, and P. Kumar (2000), A catchment-based approach to modeling land surface processes in a general circulation model 1. Model structure, *J. Geophys. Res.,* 105, 24809-24822.

Kroll, J. H., N. L. Ng, S. M. Murphy, R. C. Flagan, and J. H. Seinfeld, (2006), Secondary organic aerosol formation from isoprene photooxidation, *Env. Sci. & Tech.,* 40, 1869-1877.

Kucharik, C. J., et al. (2006), A multiyear evaluation of a Dynamic Global Vegetation Model at three AmeriFlux forest sites: Vegetation structure, phenology, soil temperature, and CO_2 and H_2O vapor exchange, *Ecol. Modell.,* 196, 1-31.

Kumar, S., C. D. Peters-Lidard, J.L. Eastman, and W. Tao (2008), An Integrated High Resolution Hydrometeorological Modeling Testbed using LIS and WRF, *Env. Modell. Soft.,* 23(2), 169-181, DOI= http://dx.doi.org/10.1016/j.envsoft.2007.05.012.

Lathière J., D. A. Hauglustaine, N. De Noblet-Ducoudré, G. Krinner, G. A. Folberth (2005), Past and future changes in biogenic volatile organic compound emissions simulated with a global dynamic vegetation model, *Geophys. Res. Lett.,* 32, L20818, doi:10.1029/2005GL024164.

Lawrence, P.J. and T. N. Chase (2007), Representing a new MODIS consistent land surface in the Community Land Model (CLM 3.0), *J. Geophys. Res.,* 112, doi:10.1029/2006JG000168.

Lehning, A., et al. (2001), Modeling of annual variations of oak (Quercus robur L.) isoprene synthase activity to predict isoprene emission rates, *J. Geophys. Res.,* 106, 3157-3166.

LeMone, M.A., et al. (2007), NCAR/CU Surface, Soil, and Vegetation Observations during the International H2O Project 2002 Field Campaign. *Bull. Amer. Meteor. Soc.,* 88, 65-81.

Leplastrier, M., et al. (2002), Exploring the relationship between complexity and performance in a land surface model using the multicriteria method, *J. Geophys. Res.,* doi:10.1029/2001JD000931.

Leuschner, C., et al. (2006), Variation in leaf area index and stand leaf mass of European beech across gradients of soil acidity and precipitation, *Plant Ecol.,* 186, 247-258.

Levis S., G. B. Bonan, M. Vertenstein, and K. W. Oleson, (2004), The Community Land Model's dynamic global vegetation model (CLM-DGVM): Technical description and user's guide. NCAR Tech. Note TN-459+IA, 50 pp.

Levis, S., C. Wiedinmyer, G. B. Bonan, and A. Guenther (2003), Simulating biogenic volatile organic compound emissions in the Community Climate System Model. *J. Geophys. Res.,* 108, doi:10.1029/2002JD003203.

Liang, X., Z. Xie, and M. Huang (2003), A new parameterization for surface and groundwater interactions and its impact on water budgets with the variable infiltration capacity (VIC) land surface model, *J. Geophys. Res.,* 108 (D16), 8613, doi:10.1029/2002JD003090.

Lyon, S.W., F. Dominguez, D.J. Gochis, N.A. Brunsell, C.L. Castro, F.K. Chow, Y. Fan, D. Fuka, Y. Hong, P.A. Kucera, S.W. Nesbitt, N. Salzmann, J. Schmidli, P.K. Snyder, A.J. TeuLing, T.E. Twine, S. Levis, J.D. Lundquist, G.D. Salvucci, A.M. Sealy, and M.T. Walter, 2008: Coupling Terrestrial and Atmospheric Water Dynamics to Improve Prediction in a Changing Environment, *Bull. Amer. Meteor. Soc.,* 89, 1275–1279.

Manabe, S. (1969), Climate and ocean circulation. I. The atmospheric circulation and the hydrology of the earth's surface, *Mon. Wea. Rev*., 97, 739-774.

Maxwell, R. M., and Miller, N. L. (2005), Development of a coupled land surface and groundwater model, *J. Hydrometerol.,* 6(3) 233-247.

Mesinger, F., G. DiMego, E. Kalnay, K. Mitchell, P. C. Shafran, W. Ebisuzaki, D. Jovic, J. Woollen, E. Rogers, E. H. Berbery, M. B. Ek, Y. Fan, R. Grumbine, W.

Higgins, H. Li, Y. Lin, G. Manikin, D. Parrish, and W. Shi (2006), North American regional reanalysis, *Bull. Am. Meteorol. Soc.,* 87, 343-360.

Mitchell, K. E., et al. (2004), The multi-institution North American Land Data Assimilation System (NLDAS): Utilizing multiple GCIP products and partners in a continental distributed hydrological modeling system, *J. Geophys. Res.,* 109, D07S90, doi:10.1029/2003JD003823.

Mo, K.C. (2008), Model-based drought indices over the United States, *J. Hydrometereol.,* 9(6), 1212-1230.

Monson, R. K., et al. (1994), Environmental and developmental controls over the seasonal pattern of isoprene emission from aspen leaves, *Oecologia,* 99, 260-270.

Naik V., C. Delire, D. J. Wuebbles (2004), Sensitivity of global biogenic isoprenoid emissions to climate variability and atmospheric CO_2, *J. Geophys. Res.,* 109, D06301, doi:10.1029/2003JD004236.

Nijseen, B. and L. A. Bastidas, (2005), Land-Atmosphere models for water and energy cycle studies, in Encyclopedia of Hydrological Sciences, vol. 5, part 17, edited by M. G. Anderson, chap. 201, pg. 3089-3102, John Wiley, Hoboken, N.J.

Niu, G.-Y., Z.-L. Yang, R.E. Dickinson, and L.E. Gulden (2005), A simple TOPMODEL-based runoff parameterization (SIMTOP) for use in GCMs, *J. Geophys. Res.,* 110, D21106, doi:10.1029/2005JD006111.

Niu G.-Y., Z.-L. Yang, R. E. Dickinson, L. E. Gulden, H. Su (2007), Development of a simple groundwater model for use in climate models and evaluation with Gravity Recovery and Climate Experiment data, *J. Geophys. Res.,* 112, D07103, doi:10.1029/2006JD007522.

Novakov, T., and J. E. Penner (1993), Large contribution of organic aerosols to cloud-condensation-nuclei concentrations, *Nature,* 365, 823-826.

Oleson, K.W., Y. Dai, G. Bonan, M. Bosilovich, R. Dickinson, P. Dirmeyer, F. Hoffman, P. Houser, S. Levis, G.-Y. Niu, P. Thornton, M. Vertenstein, Z.-L. Yang, and X. Zeng (2004), Technical Description of the Community Land Model, National Center for Atmospheric Research, Boulder, Colorado, 174 p., Accessed at http://www.cgd.ucar.edu/tss/clm/distribution/clm3.0/TechNote/CLM_Tech_Note.pdf.

Palmer P. I., et al. (2006), Quantifying the seasonal and interannual variability of North American isoprene emissions using satellite observations of the formaldehyde column, *J. Geophys. Res.,* 111, D12315, doi:10.1029/2005JD006689.

Pegoraro, E., et al. (2005), The effect of elevated atmospheric CO2 and drought on sources and sinks of isoprene in a temperate and tropical rainforest mesocosm, *Glob. Change Biol.,* 11, 1234-1246.

Pitman A.J. (2003), The evolution of, and revolution in, land surface schemes designed for climate models, *Internat. J. Climatol.,* 23(5), 479-510.

Plaza J., L. Núñez, M. Pujadas, R. Pérez-Pastor, V. Bermejo, S. García-Alonso, S. Elvira (2005), Field monoterpene emission of Mediterranean oak (*Quercus ilex)* in the central Iberian Peninsula measured by enclosure and micrometeorological techniques: Observation of drought stress effect, *J. Geophys. Res.,* 110, D03303, doi:10.1029/2004JD005168.

Poisson, N., et al. (2000), Impact of non-methane hydrocarbons on tropospheric chemistry and the oxidizing power of the global troposphere: 3-dimensional modelling results, *J. Atm. Chem.,* 36, 157-230.

Possell, M., et al. (2004), Interactive effects of elevated CO2 and soil fertility on isoprene emissions from Quercus robur, *Glob. Change Biol.,* 10, 1835-1843.

Randall, D.A., R.A. Wood, S. Bony, R. Colman, T. Fichefet, J. Fyfe, V. Kattsov, A. Pitman, J. Shukla, J. Srinivasan, R.J. Stouffer, A. Sumi and K.E. Taylor (2007), Climate Models and Their Evaluation. In: Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change [Solomon, S., D. Qin, M. Manning, Z. Chen, M. Marquis, K.B. Averyt, M.Tignor and H.L. Miller (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Robock, A., et al. (2000), The Global Soil Moisture Data Bank, *Bull. Amer. Meteor. Soc.*, 81, 1281-1299.

Rodell, M., and J. Famiglietti (2001), An analysis of terrestrial water storage variations in Illinois with implications for the Gravity Recovery and Climate Experiment (GRACE), *Water Resour. Res.*, 37(5), 1327-1339.

Rosenstiel, T. N., et al. (2003), Increased CO_2 uncouples growth from isoprene emission in an agriforest ecosystem, *Nature,* 421, 256-259.

Rosero E., Z.-L. Yang, L.E., Gulden, G.-Y. Niu, and D. J. Gochis, (2009), Evaluating enhanced hydrological representations in Noah-LSM over transition zones: Implications for model development, *J. Hydromet.*, 10(3),600-622. doi: 10.1175/2009JHM1029.1.

Saltelli, A., Chan, K., Scott, E.M. (2000), Sensitivity Analysis, John Wiley & Sons Ltd., West Sussex PO19 8SQ, England.

Schuh, G., A. C. Heiden, T. Hoffmann, J. Kahl, P. Rockel, J. Rudolph, and J. Wildt, (1997), Emissions of volatile organic compounds from sunflower and beech: Dependence on temperature and light intensity. *J. Atm. Chem.*, 27, 291-318.

Sen, O. L., et al. (2001), Impact of field-calibrated vegetation parameters on GCM climate simulations, *Quart. J. Royal Met. Soc.,* 127, 1199-1223.

Shuttleworth, W. J. (2007), Putting the "vap" into evaporation, *Hydrol. Earth Syst. Sci.,* 11, 210-244.

Simpson, D., W. Winiwarter, G. Borjesson, S. Cinderby, A. Ferreiro, A. Guenther, C. N. Hewitt, R. Janson, M. A. K. Khalil,S. Owen, T. E. Pierce, H. Puxbaum, M. Shearer, U. Skiba, R. Steinbrecher, L. Tarrason, and M. G. Oquist (1999), Inventorying emissions from nature in Europe, *J. Geophys. Res,* 104(D7), 8113-8152.

Skamarock, W. C., et al. (2005), A description of the Advanced Research WRF version 2. Accessed October 19, 2005 at http://www.wrf-model.org/wrfadmin/docs/arw_v2.pdf.

Smiatek, G., and M. Bogacki (2005), Uncertainty assessment of potential biogenic volatile organic compound emissions from forests with the Monte Carlo method: Case study for an episode from 1 to 10 July 2000 in Poland, *J. Geophys. Res.,* 110, D23304, doi:10.1029/2004JD005685.

Spatial Climate Analysis Service, Oregon State University. 1971-2000 mean monthly maximum and minimum temperature grids. Accessed June 15, 2005 at http://www.ocs.oregonstate.edu/prism/.

Stahl, C., and McElvaney, R. (2003), The Trees of Texas: An Easy Guide to Leaf Identification, 288 pp. Texas A & M University Press, College Station, Texas. pp. 1-288.

Staudt, M., et al. (2004), Isoprenoid emissions of Quercus spp. (Q. suber and Q.ilex) in mixed stands contrasting in interspecific genetic introgression, *New Phytologist,* 163, 573-584.

Stöckli, R., D. M. Lawrence, G.-Y. Niu, K. W. Oleson, P. E. Thornton, Z.-L. Yang, G. B. Bonan, A. S. Denning, and S. W. Running (2008), Use of FLUXNET in the Community Land Model development, *J. Geophys. Res.,* 113, G01025, doi:10.1029/2007JG000562

Talagrand O., R. Vautard, and B. Strauss (1997), Evaluation of probabilistic prediction systems. Proc. ECMWF Workshop on Predictibility, Reading, United Kingdom, ECMWF, 1-25.

Tao Z., A. K. Jain (2005), Modeling of global biogenic emissions for key indirect greenhouse gases and their response to atmospheric CO_2 increases and changes in land cover and climate, *J. Geophys. Res.,* 110, D21309, doi:10.1029/2005JD005874.

Tao, Z., S. M. Larson, D. J. Wuebbles, A. Williams, and M. Caughey (2003), A summer simulation of biogenic contributions to ground-level ozone over the continental United States, *J. Geophys. Res.,* 108(D14), 4404, doi:10.1029/2002JD002945.

Taylor, K.E., 2000: Summarizing multiple aspects of model performance in a single diagram, *J. Geophys. Res.,* 106, 7183-7192.

Tian, Y., R. E. Dickinson, L. Zhou, and M. Shaikh (2004), Impact of new land boundary conditions from Moderate Resolution Imaging Spectroradiometer (MODIS) data on the climatology of land surface variables, *J. Geophys. Res.,* 109, D20115, doi:10.1029/2003JD004499.

Vrugt, J. A., H. V. Gupta, L. A. Bastidas, W. Bouten, and S. Sorooshian (2003), Effective and efficient algorithm for multiobjective optimization of hydrologic models, *Water Resour. Res.,* 39(8), 1214, doi:10.1029/2002WR001746.

Wagener, T., and H. V. Gupta (2005), Model identification for hydrological forecasting under uncertainty, *Stochastic Env. Res. Risk Assess.,* 19, 378-387.

Wang X. M., G. Carmichael, D. L. Chen, Y. H. Tang, and T. J. Wang (2005), Impacts of different emission sources on air quality during March 2001 in the Pearl River Delta (PRD) region, *Atm. Env.,* 39, 5227-5241.

Wang, K. Y., and D. E. Shallcross (2000), Modelling terrestrial biogenic isoprene fluxes and their potential impact on global chemical species using a coupled LSM-CTM model, *Atm. Env.,* 34, 2909-2925.

Wang, Q., et al. (2005), On the relationship of NDVI with leaf area index in a deciduous forest site, *Rem. Sens. Env.,* 94, 244-255.

Wiedinmyer, C., A. Guenther, M. Estes, I. W. Strange, G. Yarwood, and D. T. Allen (2001), A land use database and examples of biogenic isoprene emission estimates for the state of Texas, USA, *Atm. Env.,* 35, 6465-6477.

Wiedinmyer, C., et al. (2000), Biogenic hydrocarbon emission estimates for North Central Texas, *Atm. Env.,* 34, 3419-3435.

Wiedinmyer, C., et al. (2001a), Measurement and analysis of atmospheric concentrations of isoprene and its reaction products in central Texas, *Atm. Env.,* 35, 1001-1013.

Wiedinmyer, C., et al. (2001b), A land use database and examples of biogenic isoprene emission estimates for the state of Texas, USA, *Atm. Env.,* 35, 6465-6477.

Willmott, C. J. (1982), Some comments on the evaluation of model performance, *Bull. Amer. Meteor. Soc.*, 63, 1309-1313.

Yarwood, G., Wilson, G., Emery, C., Guenther, A., (1999), Development of GloBEIS-a state of the science biogenic emissions modeling system, Final Report, Prepared for Texas Natural Resource Conservation Commission, Provided by the author.

Yeh, P. J.-F. and Eltahir, E.A.B. (2005), Representation of water table dynamics in a land surface scheme. Part I: Model development, *J. Clim.,* 18 (12): 1861-1880.

Yeh, P.J.-F., M. Irizarry, E.A.B. Eltahir (1998), Hydroclimatology of Illinois: A comparison of monthly evaporation estimates based on atmospheric water balance and soil water balance, *J. Geophys. Res.,* 103, 19,823-19,837.

York, J. P., M. Person, W.J. Gutowski, and T. C. Winter (2002), Putting aquifers into atmospheric simulation models: An example from the Mill Creek Watershed, northeastern Kansas, *Adv. Wat. Res.,* 25, 221-238.

Zhang, X. S., Y. J. Mu, W. Z. Song, and Y. H. Zhuang, (2000), Seasonal variations of isoprene emissions from deciduous trees, *Atm. Env.,* 34, 3027-3032.

Vita

Lindsey Elizabeth Gulden was born in Columbus, Indiana. She earned an A.B. in Computer Science from Harvard University in 2000. Until her matriculation at the University of Texas at Austin in fall 2004, she worked as a proofreader and copyeditor at the New England Journal of Medicine in Waltham, Massachusetts. She has accepted a research position in the energy industry.

Permanent address: 4900 E. Oltorf St. #1034, Austin, TX 78741

This dissertation was typed by the author.